Memory for temporally nonadjacent tonal centers mediated by musically salient features

Memory for temporally nonadjacent tonal centers mediated by musically salient features

By Joanna Spyra, B.Mus., M.Mus., B.A., M.Sc

A Thesis Submitted to the School of Graduate Studies in the Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

McMaster University

Doctor of Philosophy (2022)

Hamilton, Ontario (Department of Psychology, Neuroscience, and Behaviour)

TITLE: Memory for temporally nonadjacent tonal centers mediated by musically salient features

AUTHOR: Joanna Spyra (McMaster University)

SUPERVISOR: Dr. Matthew Woolhouse

NUMBER OF PAGES: xv, 111

Lay Abstract

Memory for music is often celebrated for its longevity. Music is a complex stimulus, however, and not all of its characteristics are remembered equally well. Past research has found that participants were not able to remember musical keys after a surprisingly short period of time: Farbood (2016) and Woolhouse et al. (2016) found that harmonic memory—i.e., memory for a key—lasts up to 21 seconds after a key change. Compared to nursery rhymes remembered from childhood bedtimes, this is remarkably limited. Yet this research did not fully explore which musical characteristics affect harmonic memory as it was done using simple musical stimuli: compositions made of blocks of chords. Whereas a string of chords might sound pleasant, it may not be representative of the type of music that people listen to regularly (with complex melodies and instrumentation). The focus of this project was to explore musical factors, such as melodies or rhythms, and measure how they interact with musical memory. Observing specific aspects of the stimulus gives us a window into the complexities of human memory, particularly that of the auditory domain.

Chapter 1 provides an overview of memory literature with a focus on common memory models and the musical research that supports them or contributes to their development. Here, I propose a cognitive system which integrates prominent models that otherwise describe different stages of processing complex auditory stimuli. Chapter 2 presents a detailed account of background empirical literature. This provides a basis for a series of experiments outlined in Chapters 3 and 4. These experiments investigate how components of music influence harmonic memory. Components include *Surface Features*, or ornamentations in music such as melodies or rhythms, and *Harmony*, the structure of the key itself which can make an excerpt sound more, or less, familiar. Results suggest that memory is significantly enhanced and prolonged by the addition of surface features. Furthermore, harmony that most resembles culturally familiar compositional practices also provides a memory boost when compared to random or somewhat ambiguous sequences. In Chapter 5, the implications of these results are explored with regards to the general memory models discussed in Chapter 1. Results support standard models of memory and my proposed cognitive system, as demonstrated by following the processing of my experimental musical stimuli from sound to executive function. This project suggests that more complex and musically realistic stimuli produce a significant memory boost. This puts into question traditional practices in music analysis which separate surface features into hierarchically less important positions when, in fact, the musical surface may be vital to our processing of auditory stimuli.

Abstract

Research on memory often describes the remarkable longevity of music. However, memory for music is not uniform. Cook (1987) found that participants were not able to tell apart excerpts that modulated from those that did not when the excerpt was longer than 1 minute in length. This suggests that participants were no longer able to remember, and compare, musical keys after a relatively short period of time. Farbood (2016) and Woolhouse et al. (2016) further explored the limitations of memory for tonal structures finding that, in fact, harmonic memory only lasts up to 21 seconds after modulation. However, this research was done using homophonic stimuli—arpeggios or quarter-note chords—that may not be representative of the music participants would be listening to regularly. The focus of this project was to explore how the addition of certain musical features, such as melodic or rhythmic figurations, may influence harmonic memory. Observing these possible influences may provide us with insight into the processes responsible for auditory memory and how it differs from other domains, such as speech or vision.

Chapter 1 explores prominent memory literature and music cognition experiments that support, or address concerns with, common memory models. Here, I introduce a cognitive system which reconciles music research with models by memory specialists such as Baddeley and Snyder. Chapter 2 presents a detailed account of background empirical literature, including Farbood (2016) and Woolhouse et al. (2016). Though fundamental to the exploration of temporally nonadjacent harmonic memory, this research is potentially limited in its generalizability due to the homophonic nature of the stimuli. Chapter 3 explores this limitation by testing the effects of adding surface features—melodic and rhythmic components often

used for elaboration in composition—on memory for large-scale tonal structures. Results found that harmonic memory is, indeed, enhanced and prolonged by these elaborative components, lasting up to 33 seconds, well past the limit found in previous research. Farbood (2016) further claimed that harmonic memory is significantly interrupted by new, highly harmonic excerpts. However, results from Woolhouse et al. (2016),Spyra et al. (2021) and those from Chapter 3 all question this claim as they employed stimuli that was highly harmonic. Chapter 4 investigates the contradiction by testing whether functional diatonic, functional chromatic, or random sequences degraded harmonic memory for an original key. Functional diatonic intervening information resulted in *increased* harmonic memory, directly contradicting Farbood's original findings. In Chapter 5, these results are explored in terms of prominent memory models in the field of cognition, supporting standard models of memory such as that by Baddeley and Hitch (1974) or Atkinson and Shiffrin (1968), as well as my proposed cognitive system. This is further elaborated by discussing the process of undergoing a musical judgement task from perception through to decision-making. In summary, this project suggests that more generalizable stimuli containing realistic musical features produce a significant boost in harmonic memory. Furthermore, this arguably calls into question standard practices in analysis that categorize surface features as hierarchically less important than 'deeper' harmonic events, and thus, potentially less important from a cognitive perspective. Which is to say, this evidence suggests that these features may play a vital role in remembering nonadjacent harmonic structures.

Acknowledgements

There's something empowering and humbling in writing a thesis. Beliefs, assumptions, and ideas you've put together over the years are mercilessly interrogated, usually by oneself. You're tired and confused; it's like driving to the mountains. You see them growing, so close, and yet there is always much, much more to go. My gratitude to those who have helped me on my journey, therefore, is equally momentous. I am immensely grateful for my supervisor, Dr Matthew Woolhouse, who challenged me and pushed me even when I couldn't go any further. He was supportive through difficult times and gave me space to heal, to grow, and to explore on my own. Great thanks also to my graduate committee. Thank you to Dr Ranil Sonnadarra for his mentorship and for all the meetings he didn't have time for but willingly gave anyways. Thank you to Dr Daniel Goldreich who may have begun this journey feeling apart from "all that music stuff", but was irreplaceable in all things logical and mathematical, who took me into his lab for fantastic lessons, and who was brave enough to tackle the products of Neuromatch Academy. Without these three, the road would have been long, indeed. Thank you also to Dr Scott Watter. Though cut short, his contributions to my understanding of human memory were enormous. I'm grateful for his help, support, and ruthless cross-examinations of all my theoretical and methodological assumptions.

Thank you also to the undergraduate students and volunteers who have carried me through all these years, who have had to suffer through questionable timbral choices, and who were privileged to experience my growing pains as I learned how to lead effectively, communicate clearly, teach, be a mentor, and the nebula of responsibilities that come with that. Thank you particularly to Graeme Noble who joined the team for a semester and stayed for years; I was lucky to have your hard work, constant favours, and kind friendship. My fellow graduate students, Konrad Swierczek and Maya Flannery, were a constant rock on which I could rely even at the end of it all when I disappeared often and pestered them with questions otherwise. They were, and will always be, indispensable friends. Lastly, thank you to Dr Jennifer Ostovich who was there to guide me in all things since the very beginning, and to Hazel whose beautiful face was the source of many smiles.

Finally, my undying gratitude to my family. To my partner, Dr Dobromir Dotov, who dragged me back to life, who sat through frustrations, panic, and machine learning nightmares, thank you with all my heart. To my sister-in-arms of 29 years and counting, Victoria Zyman: thank you for your wisdom, intuition, creativity, understanding, and unending support. To my parents, my beautiful niece, and my sister-in-law, thank you for supporting me through this journey and for putting up with my increasingly unrelentless work habits. And lastly, to the greatest mentor, teacher, councillor, troublemaker, enthusiastic musical omnivore, who guided me through life, who first taught me to code, who introduced me—to our parents' dismay—to the magical world of video games and all things awesome, my brother Bart Spyra, thank you. The world lost something precious when it lost you.

In memory of Bartholomew Spyra (1975-2020).

Contents

List of Figures

List of Tables

Chapter 1

Introduction

> The organization of memory and the limits of our ability to remember have a profound effect on how we perceive patterns of events and boundaries in time. . . It also allows us to comprehend time sequences of events in their totality, and to have expectations about what will happen next. Thus, in music that has communication as its goal, the structure of the music must take into consideration the structure of memory—even if we want to work against that structure. (Snyder, 2000, p. 3)

Often, memory in music is spoken of in terms of its remarkable longevity. The reminiscence bump, for example, is the tendency to remember music from times of change in one's life—moving out to university or college, getting married, etc.—better than from any other point (Krumhansl, 2017; Krumhansl & Zupnick, 2013). This suggests that music is an integral part of long-term remembering and, of course, everyday life. Music is also astonishing in patients with neural damage or degeneration. Patients with Alzheimer's disease, for example, have been shown to remember pieces of music they have learned premorbidly (before the disease), and have even been shown to successfully learn—and remember—new pieces from sheet music (Cowles et al., 2003).

The story of Clive Wearing is particularly meaningful. Clive, a professional conductor and pianist, had developed retrograde and anterograde amnesia; he had forgotten most of his life and could not retain a new memory. His wife could re-enter the room after mere moments and Clive would greet her as if it's been days. But when he played the piano, he was able to maintain memory for the entirety of the piece. In other words, musical emotion, knowledge, understanding and memory can last long after other forms of memory have degraded (Sacks, 2007, p. 337).

However, short-term memory for music is, perhaps, less studied. Paradoxically, memory for the large-scale tonal structure, for example, spans for less than a minute in duration (Cook, 1987; Farbood, 2016; Woolhouse et al., 2016). What makes music so versatile; how can memory processes be so robust and yet, so fragile?

1.1 Memory Models

Although we often talk about memory as one entity, there are many processes involved, each with their own unique characteristics and limitations. The behavioural study of memory began with Hermann Ebbinghaus (1850-1909) who gained fame for studying forgetting through memory curve experiments (see, for example, Murre & Dros, 2015). In the process, Ebbinghaus invented the practice of testing memory with nonsense syllables: a methodology that is commonly used today to avoid influences of familiarity and the meanings attached to real words. Through the study of learning and memory processes using nonsense syllables, he discovered that memory drops significantly two days after the learning phase (Goldstein, 2011, pp. 7-9). Though quite old now, Ebbinghaus's research is still fundamental in our understanding of memory. However, his experiments did not explore or describe which memory processes were involved in the forgetting curve, instead describing memory in more general terms.

Memory was not broken into subprocesses until William James in 1890, who introduced primary and secondary memories. The term, *primary memory*, represents the fleeting amount of information available in the present but not guaranteed to be recalled later (similar to echoic or short-term memory in our current understanding, as discussed below). *Secondary memory* refers to the collection of knowledge gathered over a lifetime (i.e., long-term memory; James, 1950, ch.16). Though this view does not fully describe the processes as we understand them today, both primary and secondary memories are still key components in many common memory models.

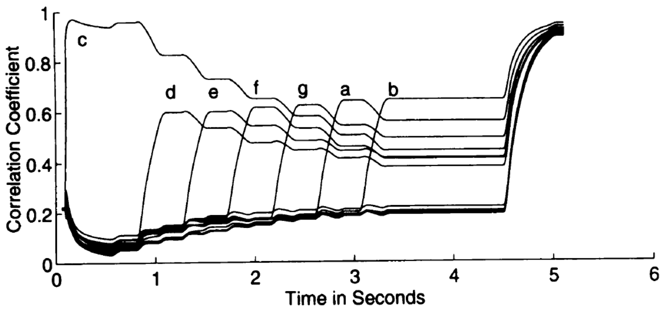
Reminiscent of James' primary memory, echoic memory (EM) is a fleeting *echo* of the surrounding environment (Snyder, 2000, p. 4). It converts perceptions into signals the brain can use for further processing. These signals are made of the most basic components of perception and are subject to millisecond decay, before any labelling or organizational processes begin (Snyder, 2000, p. 4). An example is the sound of a flute, heard by the ear but not yet identified as *flute* or even a *note* at this stage. Instead, it is perceived as a combination of frequencies, sounded at a high or low level of loudness.

The psychoacoustic model by Leman (2000) maps the process of EM from perception through to decay. Psychoacoustic models explore auditory perception through anatomical functionality. Leman (2000), for example, generates an image of neuronal firings from the peripheral auditory system and brain stem in order to create a stimulus-driven inference: a snapshot of the event in immediate memory (Leman, 2000; Leman et al., 2001). This immediate image of the auditory scene experiences rapid decay. As such, the perceptual system can experience a sort of asymmetry; when two chords are played in succession, the first chord begins to decay in memory before the second is presented. When the second chord sounds, it carries more weight in memory as it begins in full strength whereas the first chord has already decayed by some measure (see Figure 1.1). Thus, the order of presentation influences the overall perception of the event (Leman, 2000; Leman et al., 2001). EM is useful in describing music perception and is vital to early stages of auditory scene analysis (ASA) which account for Gestalt-like groupings, timbral quality, and the primary stages of stream segregation (Bregman, 1994). However, it cannot inform research on how segments of music may be compared or how judgements are made based on preceding auditory information. Thus, though a powerful demonstration of EM in music, Leman's model only describes part of the memory process.[1]

Atkinson and Shiffrin (1968) began the trend of 'standard' models (Nairne, 2002); i.e., those that categorize memory into systematically connected bins or

[1]Leman (2000) also fails to produce an image when presented with a stimulus in impoverished tones. Sine tones, for example—a timbre stripped of harmonics—provides the model with too little information to create a satisfactory echoic image (Marmel et al., 2010).

Figure 1.1
Leman's Echoic Memory Module



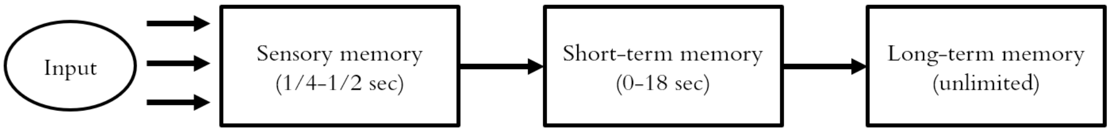
Note. Stimulus-driven inference for each note in time: each note that sounds begins with a maximal weight and begins to decay within 0.5 sec (from Leman, 2000, p. 491).

loops (see also Baddeley & Hitch, 1974). In their *modal model*, stimuli enter the memory process through *sensory memory*, the equivalent of EM. This stage acts as a buffer store for sensory information that is only active for approximately half a second. Information then makes its way into short-term memory (STM; see Figure 1.2). This component acts as an attentional control center with longer, but limited capacity. It also operates as the model's interface with long-term memory (LTM), a potentially limitless store of information gathered over a lifetime.[2] Indeed, one of the models' weaknesses is in the connection between STM and LTM. The modal model assumes that information held in STM is guaranteed to be transferred to LTM (Baddeley, 2012) which would create an incalculably vast store of information. Craik and Lockhart (1972), in particular, refuted this assumption by showing that memory is dependant on the nature of processing; deeper processing leads to better learning. Atkinson and Shiffrin also described the STM store in terms that suggest working memory processes (WM; see below) but did not separate them. This implies that a disruption to STM would automatically disrupt

[2]Much like James' secondary memory.

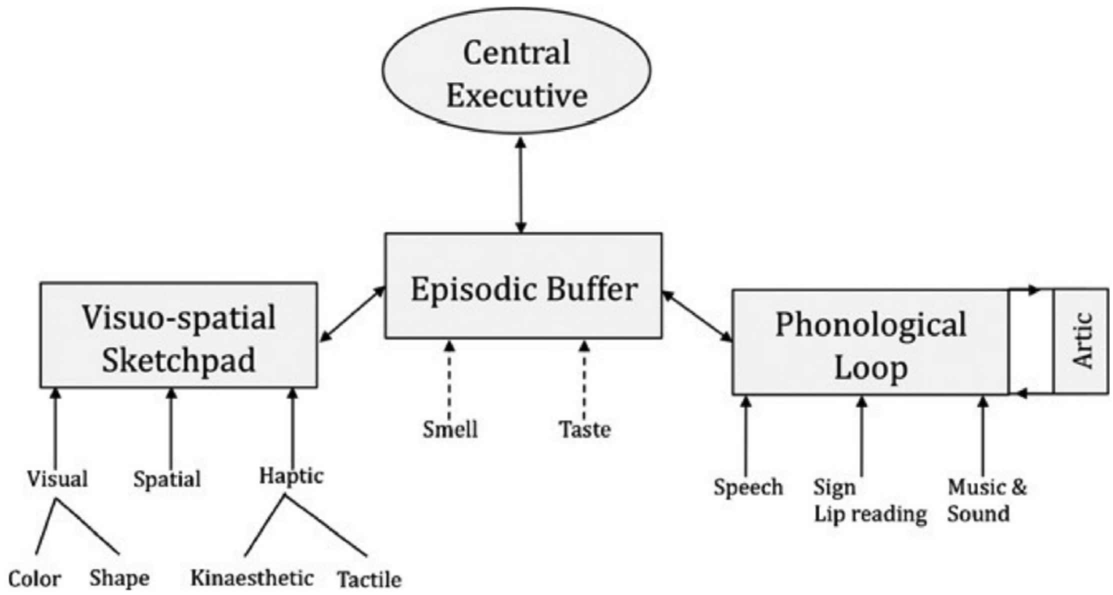
WM; a disruption amnesiac patients do not experience (Baddeley, 2012).

Figure 1.2
Atkinson and Shiffrin's Modal Model



Baddeley and Hitch's multicomponent model (Baddeley & Hitch, 1974) of WM addresses some of the concerns with the modal model by separating out and defining WM (see also Baddeley, 2003, 2012; Baddeley et al., 2010; Baddeley & Hitch, 2019; Schulze & Koelsch, 2012). Baddeley defines WM as a system with three components: the phonological loop, the visuospatial sketchpad, and the central executive/episodic buffer (Baddeley, 2003, 2012; Baddeley et al., 2010; Baddeley & Hitch, 2019). The *phonological loop* and *visuospatial sketchpad* process domain-specific perceptual information from the auditory and visual streams respectively, and maintain an activation of that information through a rehearsal process (i.e., a reinforcing *loop*; Baddeley, 2003, 2012; Baddeley & Hitch, 2019). By rehearsing, people are able to keep proceeding information active for long enough to allow processing, judgement, or action to occur. This, however, is subject to decay over time. Rehearsal loops may be further separated into subsystems, which operate at the procedural level and often reach awareness through the episodic buffer (see Figure 1.3). This suggests that music and language, for example, may be partially separated into unique subsystems with some overlapping neural resources (Baddeley & Hitch, 2019; Salamé & Baddeley, 1989). The *episodic buffer* combines information from different domain-specific loops and gates the flow of the resulting information into the central executive. The central executive works as an attentional processor of information that has been gathered and maintained in the auditory and visual loops, as well as the episodic buffer (Baddeley, 2003, 2012; Baddeley & Hitch, 2019). The central executive allows participants to make conscious, intentional decisions, and act on the stimuli being rehearsed by the other components of the model.

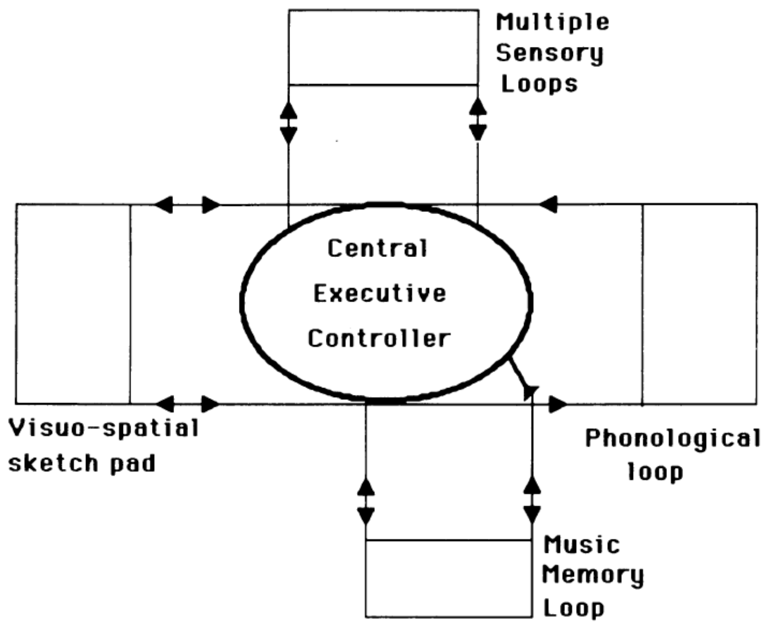
Figure 1.3
Multicomponent Model of Working Memory



Note. Updated version of the multicomponent model of working memory (from Baddeley, 2012, p. 23; Baddeley and Hitch, 2019, p. 101)

Despite the apparent success of the foregoing model, there has been debate in music cognition circles as to whether music is processed as part of the phonological loop (Lee et al., 2007) or if language and music are, instead, mutually exclusive (Berz, 1995; Deutsch, 1970; Pechmann & Mohr, 1992; Schulze & Koelsch, 2012). Berz (1995) argues that the multicomponent model does not satisfactorily account for musical memory. The brain's capacity for storing musical information in WM is well beyond that of non-musical information (spanning some 180 seconds rather than the 2 second limit that linguistic WM seems to have; Berz, 1995). Spoken numbers also seem to have no disruptive effect on tonal recall. Lastly, skilled musicians were able to chunk melodies more efficiently than novices, all of which suggests that there is a distinct musical processing component in WM (see Figure 1.4; Berz, 1995).

Figure 1.4
Model of Working Memory with Separate Tonal Loop



Note. Theoretical model of WM separating music into its own loop (from Berz, 1995, p. 362).

Deutsch (1970) also found that concurrent speech did not affect memory for a tonal pitch. Pechmann and Mohr (1992) found similar results, though more strongly pronounced in musicians than non-musicians, suggesting that this ability is further developed with musical skill. Fiveash and Pammer (2014) found that syntactic memory for sentences was significantly reduced by syntactic errors in a concurrent musical stimulus (i.e., the music had an out-of-key chord) though whether linguistic syntax effected musical memory was not tested. Finally, Schulze and Koelsch (2012) found neurological data that showed WM resources were shared between music and language; however, both had their own separate processes as well (see also Maess et al., 2001). These results together suggest that tonal WM, as a whole, is separate from the phonological loop and yet, syntactic WM resources may be shared between music and language.

The exact distinction between STM and WM is a matter of operationalization amongst different authors (Cowan, 2008). For many, such as Snyder (2000), WM is a small part of STM. For Baddeley and Hitch (2019), however, STM processes are part of a greater WM model. In essence, though, WM can be described much like Baddeley's central executive: both are responsible for attentional processes as well as executive functions which oversee computations and decision-making (Cowan, 2008; Cowan et al., 2014). In fact, evidence suggests the attentional system correlates with individual aptitudes, fluid intelligence and abilities in WM tasks (Conway et al., 2003; Engle & Kane, 2004; Kane & Engle, 2002; Redick & Engle, 2011; Unsworth & Engle, 2007). STM, on the other hand, is domain-specific and is involved in maintaining stimulus information for a short period of time, i.e., a STM task requires simple storage (e.g., a word span task) whereas a WM task requires both storage and additional processing or decision-making (e.g., a span task with concurrent prediction or calculation tasks; Conway et al., 2003).

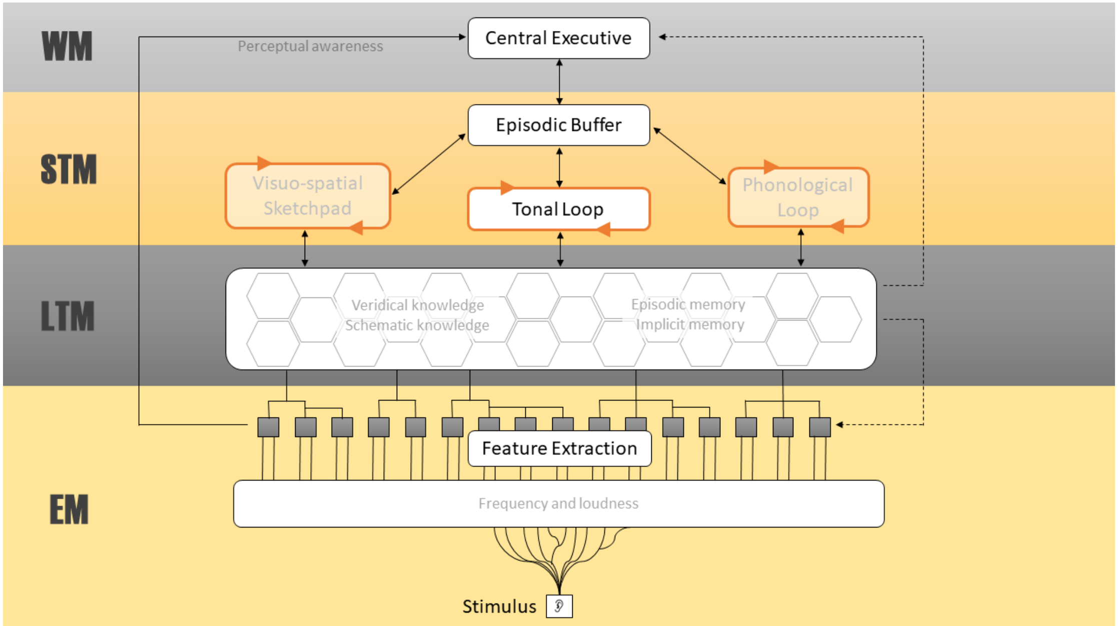
1.2 Memory in Music

Standard models are useful in describing different facets of the memory process, yet a common argument against them is that they each miss a key element that would describe all available data. In response, in Figure 1.5 I combine models based on the evidence and arguments gathered above. In this model, input enters through the sensory organ of the ear and travels as raw signals through EM creating an early image of the perceived sonic object (Leman, 2000; Leman et al., 2001). This low-level memory process precedes categorization or recognition (Snyder, 2000, p. 4). Using Gestalt-like principals, the raw input is parsed into a series of information-rich convergences, bundled into coherent events, which enable feature extraction such as instrumentation and pitch information (Snyder, 2000, p. 7) akin to the initial, pre-attentive, stages of auditory scene analysis (see, for example, Bendixen, 2014; Bregman, 1994). For example, frequencies sounding together can be grouped together at this stage to form a timbre, the particular sound or "colour" of an instrument (Alain & Bernstein, 2008).

Feature extraction connects directly to LTM, a vast store of veridical and schematic knowledge. Exemplars that are known to an individual—whether they be a favourite piece of music, painting, or novel—make up *veridical knowledge*, a term that also tends to refer to explicit knowledge (Bharucha, 1987; Guo & Koelsch, 2016). In contrast, schemata are abstracted prototypical forms derived from repeated exposure to specific types of stimuli (Agres, 2019; Bey & McAdams, 2002). In turn, these abstracted schemata implicitly guide the processing of incoming information, perceptions, and actions. Schemata enable us to make sense of and categorize musical stimuli irrespective of whether we have veridical knowledge of it (Bharucha, 1987; Guo & Koelsch, 2016). Schemas compiled over a lifetime through experience and enculturation, as well as those created from preceding acoustic data, help interpret the scene, separate sources of acoustic information,

Figure 1.5

Proposed Model of Memory with Focus on Musical Processing



Note. My proposed standard-type memory model; this model takes its organization and the features of EM from Snyder (2000), combines it with the multicomponent model of WM (Baddeley & Hitch, 2019), and is informed by music cognition experiments to include a tonal loop (Berz, 1995; Deutsch, 1970; Pechmann & Mohr, 1992).

and influence the way we perceive a stimulus by strengthening and creating expectations of how it might continue (Alain & Bernstein, 2008; Bharucha, 1987, 1994; Bregman, 1994; Justus & Bharucha, 2001; Rogers & Bregman, 1993; Scheirer, 1996). This has even been demonstrated in non-musicians who have been shown to have a sophisticated understanding of musical knowledge (Cui et al., 2022). A particular feature of LTM, in general, is that it does not appear to have any storage or time constraints, unlike other levels of memory which only operate for relatively brief durations (Chen & Cowan, 2005; Cowan, 2008; Craik & Lockhart, 1972).

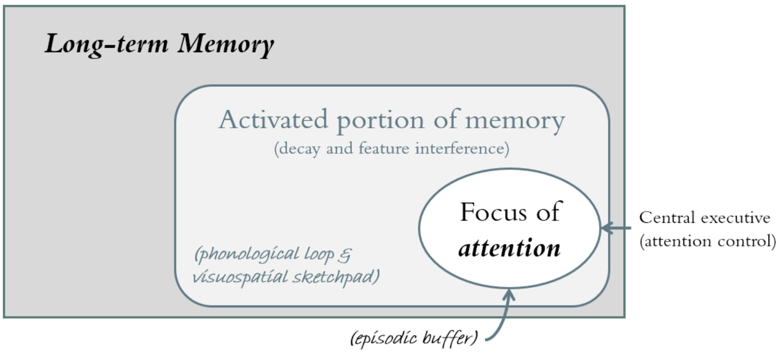
In Snyder's conception of musical memory, feature extraction and LTM both connect to STM and WM, which are responsible for the interpretation of the current experience and influence our immediate decisions (Baddeley & Hitch, 2019; Snyder, 2000, p. 6). LTM, with its schematic knowledge, can help various STM processes, e.g., maintaining an understanding of a piece of music as it meanders through various tonal centers. Though it is possible that syntactical memory somewhat overlaps with the phonological loop (Fiveash & Pammer, 2014; Schulze & Koelsch, 2012), a separate tonal loop has been specified in light of evidence from Berz (1995) and Pechmann and Mohr (1992). The tonal loop is responsible for maintaining a memory of the musical stimulus over a short period of time and decays as time increases (Berz, 1995). Due to decay and capacity limits of the loop, a more complex stimulus may require additional STM resources. In such a case, the episodic buffer may hold additional information for future processing as the tonal loop rehearses immediately preceding information (Baddeley & Hitch, 2019). Lastly, the central executive is the only center of attention in this system, as is used for bringing attention to aspects of sounds (for example, a particularly loud or strange sound) and for various judgement tasks (Conway et al., 2003; Engle & Kane, 2004; Kane & Engle, 2002; Redick & Engle, 2011; Unsworth & Engle, 2007).

Though my proposed cognitive system is an attempt at an inclusive model of musical processing in memory, it is important to note that many alternative models exist that do not conform to the typical structure of a standard model. Assumptions made by standard and non-standard models will be compared in following chapters.

1.3 Alternative Models

STM, WM and LTM are deeply interconnected, such that it can be difficult to tease them apart into separate processes. Indeed, researchers such as Cowan argue that they all activate the same memory stores (Cowan, 2008; Craik, 2020). In Cowan's model, STM is a temporarily activated subset of LTM which decays over time unless it is refreshed in some way (see Figure 1.6). WM is a further subset of STM, the focus of attention, and is limited in chunk capacity. WM storage is further divided into *central* storage, which categorizes information (i.e., it is involved in allocating resources much like Baddeley's central executive), and *peripheral* storage in which sensory information is represented in separate modalities (i.e., Baddeley's phonological loop; Cowan et al., 2014).

Figure 1.6
Cowan's Theoretical Modeling Framework



Note. Modified from Cowan (2008, p. 326).

Though Cowan's model takes steps to blend the components of memory into a more cohesive structure, it still groups processes into bins much like Baddeley did. Both are, therefore, subject to similar criticisms: that memory is based on activation, rehearsal, and is subject to decay (Nairne, 2002). Alternative approaches, such as Lewandowsky et al. (2004) and Nairne (2002), challenge the existence of these processes and, in contrast, portray memory as cue-based, where the success of recall hinges on the appropriateness of cues. What is described as memory *decay* in standard models is instead explained as *cue overload*: the point at which the cue contains more information than the target and thus, cannot easily identify the appropriate memory (Nairne, 2002; Neath & Surprenant, 2003, pp.121-138).

Lewandowsky et al. (2004) challenged the idea of time decay by presenting participants with a list of six items at 400, 800, or 1600 msec/item with or without articulatory suppression. They hypothesized that, all things being equal, should memory decay over time, the 1600 msec/item condition would experience the most deterioration. In fact, results should show a fanning in the interaction between time and speed in which faster conditions experience the least time-based forgetting, followed by 800 msec/item condition and lastly the 1600 msec/item condition. However, no evidence supporting time-based forgetting was found, only main effects of speed, which simply suggested that participants could learn to recall lists at different speeds; this on its own cannot be explained well using time-based models (see also Duncan & Lewandowsky, 2005). However, Cowan and Aubuchon (2008) question these results, arguing that there could yet be non-articulatory forms of rehearsal at work.

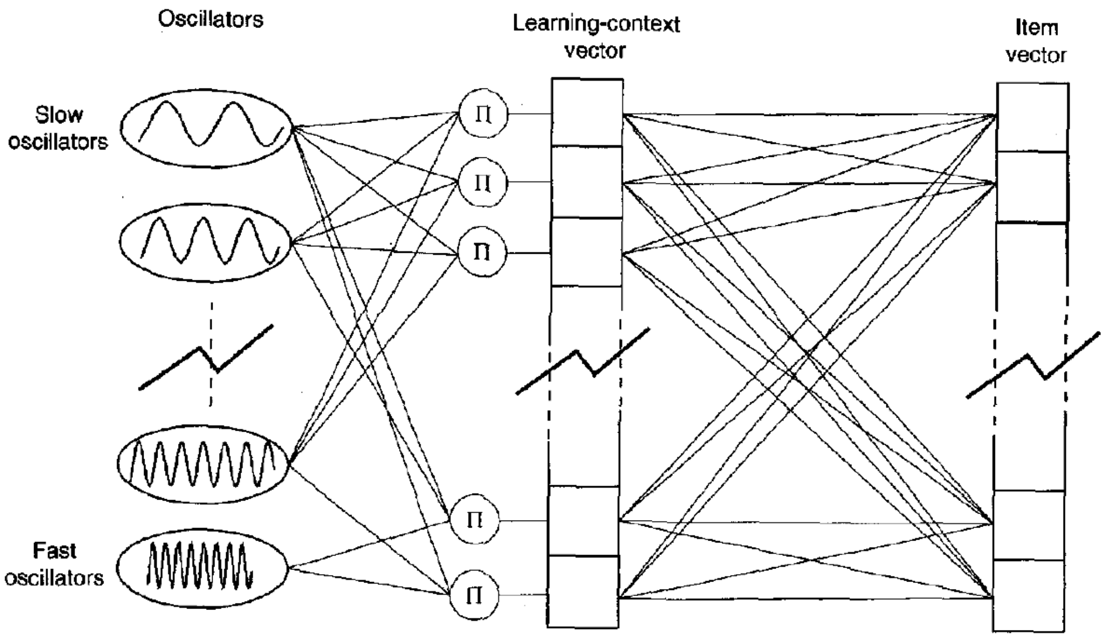
Unitary models, such as the feature model (Nairne, 1990, 2002) or the OSCAR model (Brown et al., 2000) completely reject the ideas driving standard models. They believe that there is no connection between activation strength and memory, that there is no rehearsal, and argue for cue-based forgetting instead of decay over time as described above (Nairne, 2002). Unitary models assume there are similar processes for STM and LTM, only retrieval cues differ for each.

The feature model (Nairne, 1990, 2002) argues that STM is based on cues; what "sits" in STM isn't an activated item, but an activated network of cues. In most cases, these cues are remnants of *past processing records* (previously presented items that have degraded through interference). It's not the match between cues and LTM that's important, it's how well cues uniquely specify the targeted items. Therefore, short-term forgetting occurs because cues become poor predictors, overwritten by new items as a function of similarity. Increasing similarity between stimuli tends to reduce the predictive power of common features and the cue becomes overloaded as it can predict several items at once (thus lowering the chance of remembering the correct item). Performance similarly declines as a list grows longer.

Due to cue-driven processes, the feature model can explain many things other models cannot. For example, people forget over time because retrieval cues change over time, not because the activation fades, which could explain the results found by Lewandowsky et al. (2004). Additionally, it can handle the fact that people don't forget, or that memory improves with time, provided that cues aren't interfered with, or are reinstated in some way.

The OSCillator-based Associative Recall (OSCAR) model is another unitary model of memory. This model relies on oscillator-based remembering (Brown et al., 2000). Items of a serial order are input as a vector into this model (see Figure 1.7, right side). Vectors are made of elements which create various patterns of activation based on the input (the number of elements can be modified as needed). Each element of the *item vector* connects to each element in a second, *learning-context vector* (Figure 1.7, middle section). A new item vector is formed for every item that is input into the model. Each element of the learning-context vector updates at a different frequency, determined by the product of four oscillators each. This ensures that every element of the learning-context vector is updating at a unique frequency. This is important as the patten of oscillations can provide regularities (or a complete lack of regularity if so desired).

Figure 1.7
The OSCAR Model



Note. The OSCAR model by Brown et al. (2000, p. 131).

For a more approachable understanding, think of a clock face. The hour hand represents the longest oscillator with a step size of one hour, the minute hand has a smaller step size, and the second hand is the smallest oscillator of them all. If an item is presented at 4:00, another at 4:05, another at 4:10 and so on, the unique combination of hands forms a representation of that item (Brown et al., 2000). Consequently, an item at 4:05 will also cue—in memory—items closely positioned to it on the clock: the items at 4:00 and at 4:10. Furthermore, the clocks hands will repeat certain positions every n timesteps causing a pattern reminiscent of octave-equivalence. As such, an item presented at 4:00 will be closer in memory to the item at 5:00 than to one at 4:25, even though 4:25 is "closer" in distance; the combination of oscillations resonates more closely on the hour mark, just as middle C resonates more closely with C an octave higher than with an F.

The clock is, however, only a representation of the simplest parameters of the model. OSCAR can also be made to represent free recall, for example, by employing oscillator combinations with resulting patterns that do not repeat (Brown et al., 2000, p. 136). As powerful as this model is, it does not account for evidence that supports assumptions made by standard models: memory would not decay over time with such a model, nor is there any room for rehearsal. Would recall, then, be perfect every time? It has further been argued that unitary models have difficulty accounting for interactions between duration effects and articulatory suppression (Baddeley, 2000). Given there is no room for rehearsal processes in the feature or OSCAR models, it would be logical for articulatory suppression to play little part. It has been suggested, however, that articulatory rehearsal plays a role in maintaining memory for melodies (Nees et al., 2017), a finding that unitary models would have a difficult time accounting for.

Chapter 2

The Perception of Key Relationships

> If we now consider that, in addition, the return to the tonic coincides with the formal conclusion—as it does in this consequent—and that it thus signifies a return to the harmonic point of departure, we see that the motion has reached its goal: form as well as harmony have closed their circle; and for this reason we call such a conclusion a full close, a perfect cadence. (Schenker, 1954, p. 217)

Schenker describes a common way of thinking in music analytical circles: that tonality is a cyclical,[1] holistic process that meanders through harmonic terrains, but ultimately returns home. From short phrases to entire movements, music often refers to itself across time (Laitz, 2012, pp. 371-384). Sonata Form, for example, is entirely built on this idea (Stravinsky, 1970, p. 41) where the beginning and ending of the piece reflect one another (Kostka & Payne, 2004, pp. 332-334). A

[1] Cyclical in the sense that it begins and returns to a tonal region.

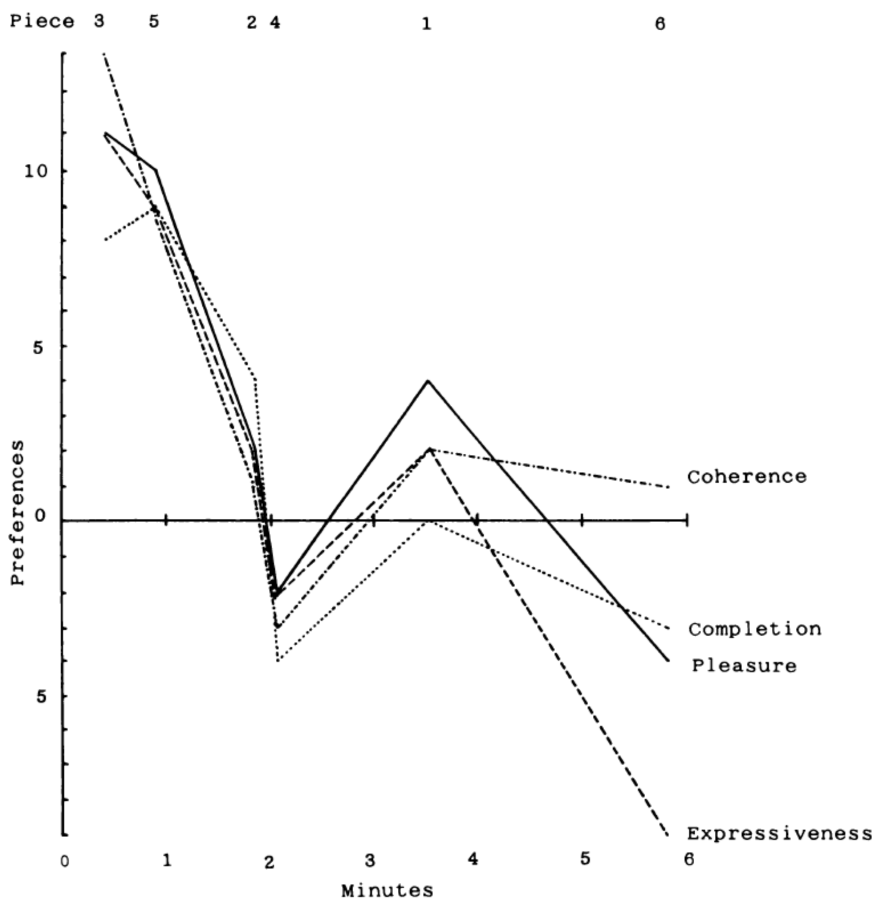
vital element of this relationship is the juxtaposition of musical keys; tonal regions are used as a tool by composers to create a feeling of familiarity upon their return (Laitz, 2012). However, little work has been done thus far to measure how well participants can perceive these tonal structures and relationships within them. In fact, it has been suggested that that there is little to no sensitivity to global harmonic structures (Granot & Jacoby, 2011).

2.0.1 Cook (1987)

Cook (1987) was one of the first researchers to study large-scale tonal structures. The term *large-scale tonal structure* is perhaps deceiving, as a tonal structure does not need to be large to challenge the limits of our perception. Cook (1987), for example, noticed that structural relationships between movements are weaker than those within a movement. But it was unclear whether large-scale tonal closure—another word for the cyclical harmonic relationship in music—affects listeners' aesthetic responses to music and to what temporal limits this effect manifests. In two experiments, participants were presented with two consecutive versions of piano pieces: one that remained in a single musical key, and one that modulated (i.e., moved to a different key). Six excerpts were manipulated in this way and ranged in duration from 30 seconds to 6 minutes. Each was played live by a professional musician. Participants were asked to make two-interval forced choice preference judgements based on characteristics of the excerpts, including expressiveness, coherence, pleasure, and sense of completion. A strong preference for either modulating or non-modulating excerpts in any of these characteristics would indicate there is an effect of tonality on aesthetic and/or structural perception.

Cook found significant differences for only the shortest excerpts (30 sec and 1 min in length), suggesting that participants could not distinguish modulating from non-modulating pieces that were over one minute in length (Figure 2.1). He did not, however, manipulate (or, indeed, measure) the exact length of modulating phrases within the excerpts providing little more than a hint of the memory limitations involved in large-scale tonal perception.

Figure 2.1
Preferences for Modulating Versions



Note. Positive results indicate a preference for the modulating version. Only the shortest two excerpts are significant (from Cook, 1987, p. 201).

Interestingly, Cook also presented participants with an excerpt which did not modulate (i.e., both presentations of this excerpt were exactly the same). In both experiments, participants preferred the second presentation of this excerpt significantly more than the first one, a result that was not found in the stimuli that modulated. This was explained as a preference for repetition though if one were to assume that participants cannot process harmonic closure for excerpts longer than one minute as the above results suggest, one would expect this same pattern in longer modulating stimuli as well; such a pattern is not, in fact, reflected in the graph in Figure 2.1, with the exception, perhaps, of Piece 6's *Expressiveness* result.

Without more detailed descriptions of Cook's stimuli, it is difficult to say what could have been the driving force in this difference. Each moment in music is influenced by preceding moments (Leman, 2000; Leman et al., 2001; Temperley, 2007, p. 89) and the fact that Cook's stimuli modulated (or did not) may itself have changed the way participants perceived these pieces. However, even this localized view—where connections are made between single events—is simplistic and does not account for the global effects[2] that are vital in sonatas and other forms of large-scale tonal structures (Tillmann & Bigand, 2004). Cook's measures did not account for either of these effects, making it impossible to say what was driving the results, how elements in his stimuli were processed, or how long the second key needed to be to override the original in memory (see Snyder, 2000, pp. 3-15, 47–53).

[2]**Global effects:** between musical structures in extended sequences; **Local effects:** between consecutive events.

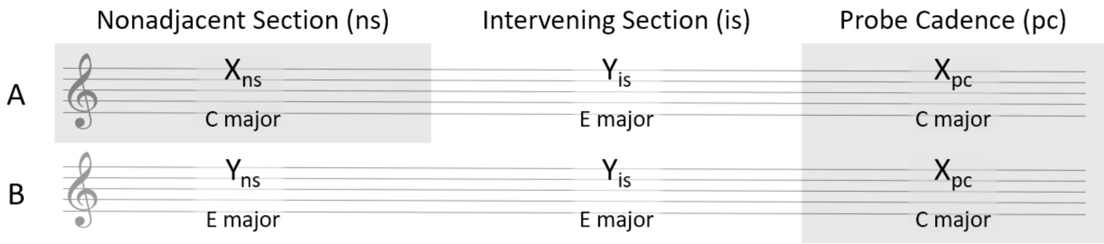
2.1 The Nonadjacent Paradigm

Such questions were addressed in experiments performed almost 30 years later. These studies employed the *nonadjacent paradigm*, a deceptively simple yet powerful tool for studying large-scale musical structures. In this paradigm, stimuli are divided into three sections: the nonadjacent section (*ns*), intervening section (*is*), and probe cadence (*pc*). Each section can vary in any musical characteristic the researcher is interested in pursuing; in the experiments that follow, for example, they will be organized by key. By varying the modulations between sections, one can target the effects harmony has on perception and memory.

Consider an experiment—a modified version of Cook (1987), for example—with two conditions: one which modulates so that the stimulus begins and ends in the same key (i.e., an X-Y-X key relationship between the sections) and one where the first two sections remain in the same key but modulates for the probe cadence (i.e., a Y-Y-X key relationship). For clarity, let us label each section by its harmonic relationship to the probe cadence—the target section in this paradigm—and subscript each appropriately: an X-Y-X relationship thus becomes X_{ns}-Y_{is}-X_{pc} and Y-Y-X becomes Y_{ns}-Y_{is}-X_{pc} (ns = nonadjacent section, is = intervening section, pc = probe cadence; see Figure 2.2). By comparing these two conditions, one can isolate the local effect of Y_{is} to X_{pc} from the global effect of X_{ns} on X_{pc}, the probe. By calculating the difference in participant responses between two stimuli that are matched in every way except for the key structure, one can test how strong the global and local effects are on harmonic memory.

Figure 2.2

Nonadjacent Key Relationship



Note. Nonadjacent key relationships as illustrated in Spyra and Woolhouse (2021, p. 4). Capital letters represent key relationships between sections and subscripts specify which section is discussed.

2.1.1 Woolhouse et al. (2016)

This paradigm can be modified in hundreds of ways, limited only by the researcher's imagination. Woolhouse et al. (2016), for example, manipulated stimuli in both duration and modulation. Harmonic relationships were as described above: in the modulating condition, key relationships were X_{ns}-Y_{is}-X_{pc}, and in the non-modulating condition, Y_{ns}-Y_{is}-X_{pc} (Figure 2.2). The intervening section varied in duration from 2-12 chords. As such, the harmonic relationships between sections could be used to observe two effects that Cook (1987) was missing: the first tested participants' ability to remember a tonal structure after it has modulated away from the original key and the second measured how long this memory could last. Participants were asked to rate the probe cadence on goodness-of-completion using a 7-point Likert-type scale (1 = no completion, 7 = high degree of completion); if participants were able to remember the key in X_{ns}, the sense of completion of the probe (X_{pc}) would be higher than in an otherwise identical Y_{ns}-Y_{is}-X_{pc} condition.

Stimuli were presented in Shepard tones (Shepard, 1964) to approximate equal overall pitch height. All stimuli began with a I-IV-V-I chord progression (the nonadjacent section) and ended on a V-I cadence (the probe). Intervening sections ranged 2.5-8.75 seconds in length and were made from 2-12 chords respectively. A repeated-measures design was used and musicians and nonmusicians scores were compared separately.

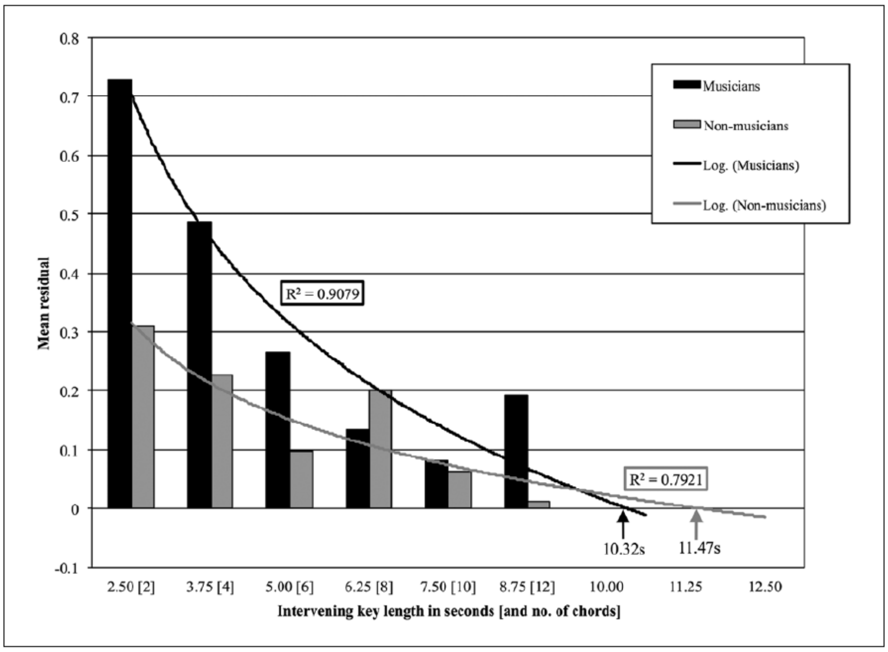
Ratings were higher for modulating stimuli, indicating that participants were, indeed, able to hold a tonal structure in memory. By calculating difference scores (termed *residuals* in Figure 2.3) between modulating and non-modulating conditions, Woolhouse et al. (2016) could pinpoint at which moment, if at all, the difference decreased to zero; this point would signify that memory for X_{ns} had decayed completely. Memory for a harmonic structure (i.e., the nonadjacent section) lasted for approximately 11 seconds (see Figure 2.3) in both groups, suggesting that memory is somewhat limited for musical keys, regardless of musical skill.

2.1.2 Farbood (2016)

Subsequently, Farbood (2016) conducted a similar series of experiments, also using the nonadjacent paradigm. Instead of Likert-type ratings, however, Farbood used tension judgements. Sharp increases in tension have been shown to correspond with tonal modulations due to their perceptual novelty (Bigand & Parncutt, 1999; Farbood, 2016; Lerdahl, 1988, 2001; Lerdahl & Krumhansl, 2007). Arpeggiated chords were used to establish ecological validity and help alleviate awkward voice leading. Unlike Woolhouse et al. (2016), a piano timbre was used. Arpeggiated chords changed every 1.5 seconds.

In Experiment 1, nonadjacent keys were always congruent (i.e., X_{ns}-Y_{is}-X_{pc} key relationship). The intervening section was modified in duration to last between

Figure 2.3
Musicians' and Nonmusicians' Retention of Key

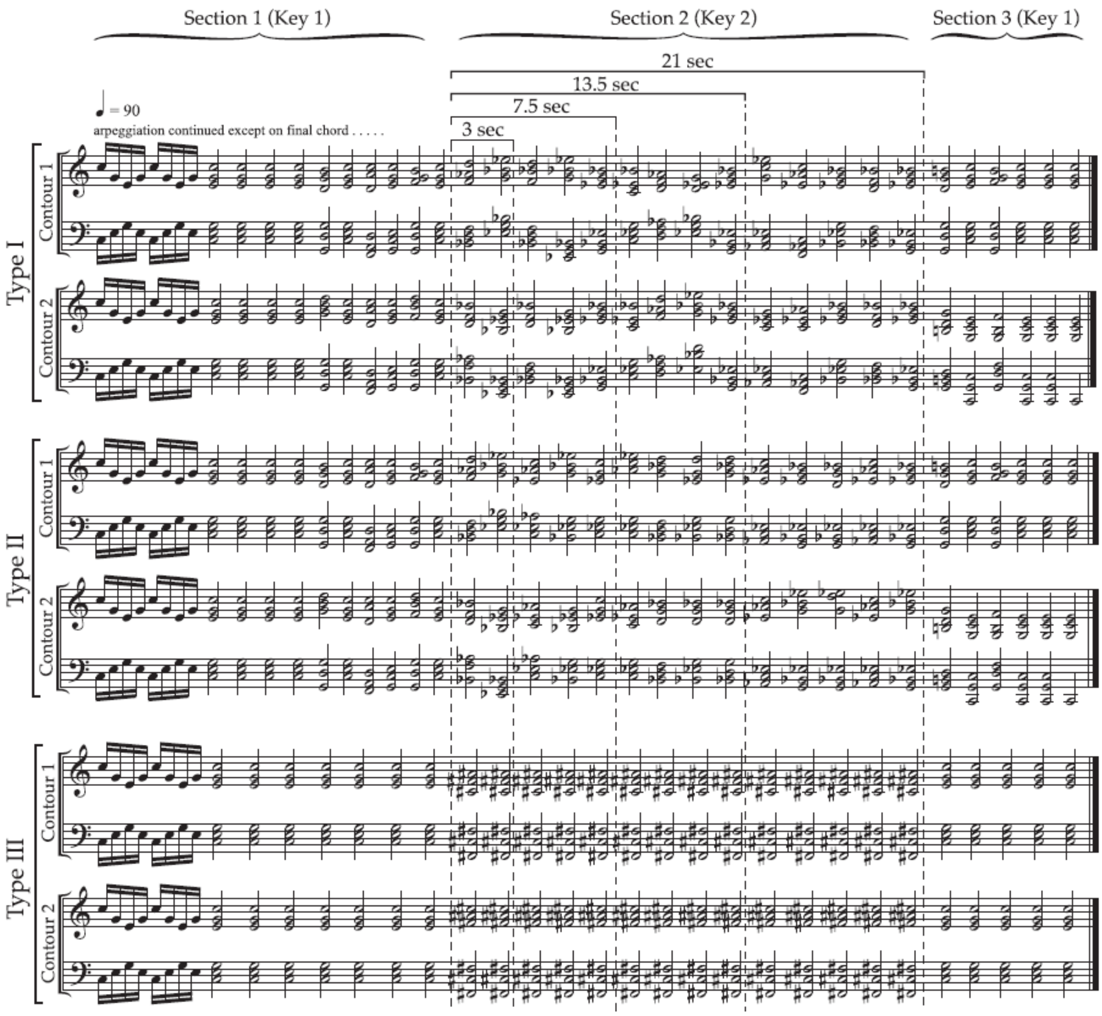


Note. From Woolhouse et al. (2016, p. 8).

0-21 seconds; the zero-second condition provided a version of stimuli that did not modulate (i.e., X_{ns}-X_{pc}). Instead of forming a baseline for local effects as Woolhouse et al. (2016) did, this forms a tension baseline for a regular, albeit short, non-modulating musical stimulus. There were three types of harmonic sequences used in Experiment 1: in Type I, the intervening section was composed in a functional Western style (i.e., it corresponded to all tonal rules of composition in traditional Western diatonic practice; see Auhagen & Vos, 2000; Laitz, 2012, p. 3); the Type II intervening section was meandering and unpredictable; and in Type III, it was made of a single repeating tonic chord a tritone away from the original tonic (see Figure 2.4). This results in three conditions with varying degrees of tonality in the intervening sections; Type I would sound most familiar to participants and had a clear sense of key (including a key-defining cadence at the end), Type II was random, and Type III would sound the most ambiguous.

Figure 2.4

Farbood's Stimuli in Three Sections and Types



Note. The nonadjacent key paradigm with X_{ns}-Y_{is}-X_{pc} structure. Y_{is} in Type I follows Western tonal compositional rules, in Type II is meandering, and in Type III is made of a single repeated chord (from Farbood, 2016, p. 76).

Tension was rated continuously on a scale of 0-100 throughout the stimulus. In analysis, slopes and magnitudes of tension were calculated for three second intervals of continuous data after each modulation in Type I and II conditions (but only 1.5 seconds for the Type III condition[3]). Tension was compared between both points of modulation: X_{ns} to Y_{is} and Y_{is} to X_{pc}. The tension slope of the second modulation was indicative of how well the first key was retained in memory: a negative slope suggested the original key was recalled as a decrease in tension meant the original key was still the primary key context. A positive slope meant the original key was forgotten and the new key had replaced it in memory.

Results show a residual effect of the original key was still present, though faint, at 21 seconds, suggesting even this duration, well past that found in Woolhouse et al. (2016), was not enough to completely erase the nonadjacent key in memory. However, this result only pertains to stimuli where the intervening section had lowest tonality: Type III stimuli extended memory for over 21 seconds while well-formed Type I (and even Type II) versions decreased memory drastically suggesting that the degree of tonality, and perhaps other characteristics of music, may have significant effects on harmonic memory (see Figure 2.5).

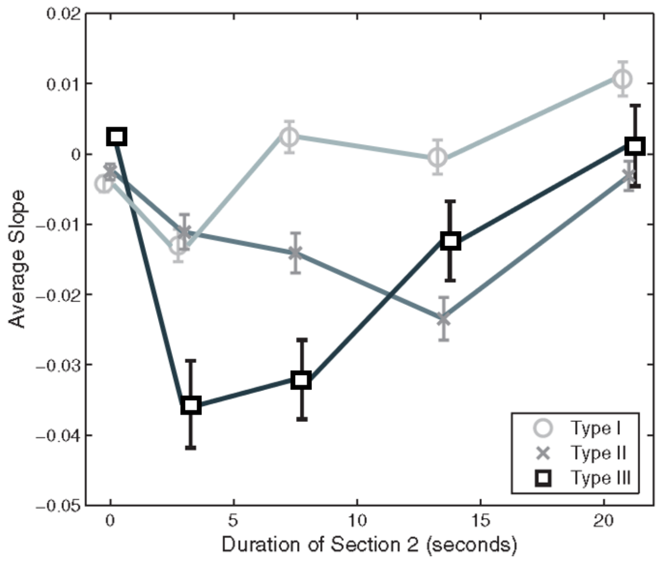
In Experiment 2, only Type III stimuli were used, ensuring that only the most successful condition from Experiment 1 was tested. These, however, were divided into three subtypes: non-modulating (which remained in a single key throughout), closed sequences (i.e., congruent key relationships), and open sequences in which each section modulated to a different key (i.e., a Z_{ns}-Y_{is}-X_{pc} key relationship). The duration of the intervening section was modified to last between 0-45 seconds.

Results found no significant differences between open and closed sequences after 10 seconds. These results would seem to support the findings in Woolhouse et al.,

[3]This was done to account for no transition chord in Type III conditions deeming a full three seconds unnecessary.

Figure 2.5

Results of Three Types from Experiment 1



Note. Type III had the most robust results, while Type I—which are similar to stimuli from Woolhouse et al. (2016)—did not reach significance (modified from Farbood, 2016, p. 81).

2016 if not for the fact that the comparison was made between unequal group sizes: there were 5 non-modulating, 30 closed, and 10 open sequences. Furthermore, comparisons were made by administering many two-tailed t-tests which increase the already heightened probability of error.

Like Woolhouse et al. (2016), the two studies by Farbood addressed the effects that Cook (1987) was missing: they reconfirmed that participants could, indeed, distinguish between modulating and non-modulating stimuli, and they tested the exact duration of harmonic memory (in this case 20 seconds). Furthermore, results from Farbood (2016) and Woolhouse et al. (2016) suggest that memory for tonal structures do decay over time, supporting standard models of memory[4]. However, both manipulated intervening section duration by adding chords. In Woolhouse et al. (2016), for example, an intervening section with duration of five seconds would always be comprised of six chords. Farbood (2016) similarly added repeated arpeggios in proportion to the length of the intervening section. This causes the number of chords to systematically vary with levels of the factor, *Duration*, creating a possible confound in the design. As such, it is unclear whether it was a decay over time or else the amount of intervening information that affected memory. Conversely, both time and number of events (e.g., chords) could have an effect on memory (Akiva-Kabiri et al., 2009).

2.1.3 Spyra et al. (2021)

Spyra et al. (2021) sought to address this issue by manipulating duration and number of events separately. Also using the nonadjacent paradigm, the intervening section in this study lasted either six or nine seconds in duration and was composed of either four or six chords, leading to the combinations found in Figure 2.6. Participants were asked to rate the degree to which the probe cadence completed the musical phrase on a Likert-type scale from 1 (not at all) to 7 (strong sense of closure). A piano timbre was used for this experiment.

Results showed a significant memory decay as time increased, but no significant effect of the number of events (chords), suggesting that it is, indeed, time that influenced memory degeneration. Figure 2.7 shows difference scores between

[4]A cue-based memory model would hypothesize that memory was overloaded with cues as time went on. However, in a stimulus consisting of a single repeated chord as in Farbood (2016), there is, presumably, no such increase of cues.

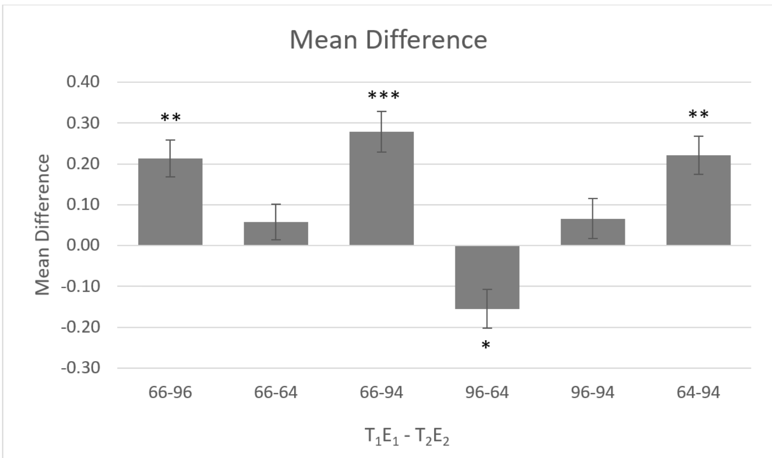
Figure 2.6

Time and Events Juxtaposed

Note. Four sample stimuli comparing intervening sections with factors *Time* (T = 6 or 9 sec) and *Events* (E = 4 or 6 chords) from Spyra et al. (2021, p. 217).

conditions. The x-axis is divided between juxtapositions of two conditions each: the first number of each pair represents the time (T) spent in the intervening key, while the second number represents the number of events (E) present in that condition. As such, the first bar on the left (66-96) is a difference taken from average ratings for a condition in which the intervening section was 6 seconds long and was composed of 6 chords from a condition that lasted 9 seconds and was composed of 6 chords. In this case, the only difference between conditions (T_6E_6-T_9E_6 as labeled in the paper) was the duration of the intervening section.

Figure 2.7
Difference Scores between Time and Events



Note. Means difference values between conditions *Time* and *Events* (from Spyra et al., 2021, p. 220).

Any time a condition lasting 6 seconds was subtracted from a 9 second condition—as in the first, third, and last bars in Figure 2.7—the difference score was significant and positive. The opposite is true of 9 seconds subtracted from 6 seconds; the results were significant but negative suggesting that as duration increased, goodness-of-completion rating decreased. However, the same cannot be said of comparisons between events as in the second and fifth bars in Figure 2.7.

The results in Spyra et al. (2021) were in line with the findings of Cook (1987), Farbood (2016), and Woolhouse et al. (2016). There was, indeed, a significant decay of memory for a nonadjacent musical key over time. This finding supports the predictions of the standard models of memory proposed in Chapter 1: despite the findings of Lewandowsky et al. (2004), there is consistent evidence suggesting that harmonic memory decays over time instead of being replaced or perturbed by new information. Furthermore, cue-based models of memory are not supported in these studies; from the perspective of these models, findings from Spyra et al. (2021) should have found opposite effects, as events would have caused the cues to represent multiple items in musical space and time decay certainly would not have affected memory in either the feature model (Nairne, 2002; Neath & Surprenant, 2003) or OSCAR (Brown et al., 2000). Results, therefore, do not support the idea that new events replace cues in memory as suggested by alternative models of memory.

Chapter 3

Surface Features

> In the language of nineteenth-century organicism, form is analogous to the surface appearance of an organism, the articulation of its limbs, whereas structure is analogous to its skeleton. Nevertheless, the two dimensions are clearly closely related, for (in a top-down view) the structure motivates the surface configurations and (in a bottom-up view) the surface generates and shapes the structure. (Jan, 2010)

In the previous chapter, I explored three studies that used nonadjacent key relationships to test harmonic memory after modulation. Spyra et al. (2021) confirmed that memory decays over time, instead of being replaced by new information. Woolhouse et al. (2016) suggested that memory for a key decays completely by 11 seconds after modulation. Finally, Farbood (2016) found that harmonic memory can be drastically affected by tonal characteristics of the musical stimulus. These studies, however, all shared a foundational limitation: the stimuli used were largely homophonic in nature, composed in quarter-note chords (Spyra et al., 2021; Woolhouse et al., 2016) or arpeggiations (Farbood, 2016). Though easy to control and harmonically well-formed, such stimuli may not generalize to music more commonly heard by participants and, thus, may be missing key elements of composition that significantly influence global harmonic memory. Elements especially lacking in homophonic stimuli are *surface features*, compositional features of the musical surface such as rhythmic or melodic elements, instrumentation, or timbre.

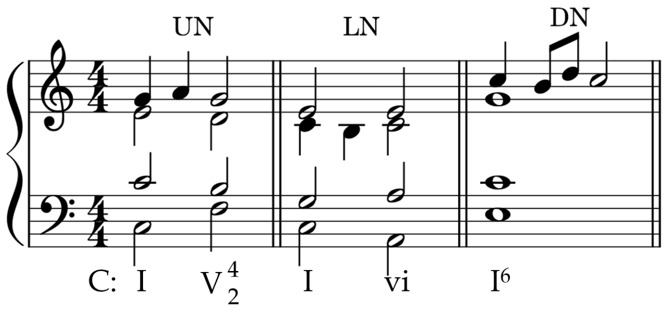
3.1 Surface Features

Common in Advanced Analysis classrooms, the methods of Heinrich Schenker (1954) teach students that in order to get to the hierarchically deepest analysis of a piece, one must remove the musical surface completely. This involves stripping down the harmonic landscape until only the deepest musical background (or *Urstaz*) remains (Bharucha, 1994; Lerdahl and Jackendoff, 1996; Pankhurst, 2008, pp. 54-55; Schenker, 1954, p. 43). The lowest hierarchical level, and the first to be eliminated, is comprised of surface features, the more elaborative characteristics of composition (i.e., melodic components, rhythmic activity, instrumentation etc.). In this chapter, I focus on melodic and rhythmic components of the musical surface.

Melodic components include musical decorations in the form of both melodic and rhythmic *figuration* (Ln. *figurare*: the shape or form). The primary use of figuration is to elaborate the harmonic progression of the composition by adding notes or by rhythmically displacing notes (Aldwell et al., 2010). Furthermore, figuration is a harmonic embellishment that can occur in any voice, or many voices simultaneously.

There are various types of *melodic figuration*. Included are passing tones, neighbour tones and chordal skips. Passing and neighbour tones are non-chord tones—tones that do not belong to the chord that is occupying that space in time—and move by one step on the staff (see Figure 3.1). As non-chord tones, these are comparatively dissonant to the surrounding harmony and resolve fairly quickly. Passing tones continue the movement of the music in the same direction, usually occurring on the offbeat (in between major beats of the measure). Neighbour tones return to the previous chord tone, thus reversing the direction of the movement. As such, there are a variety of neighbour tones: upper and lower, which refers to the direction in which they move, and double neighbour tones, i.e., elaborations comprised of two non-chord tones (see Figure 3.1 for examples).

Figure 3.1
Neighbour Tones



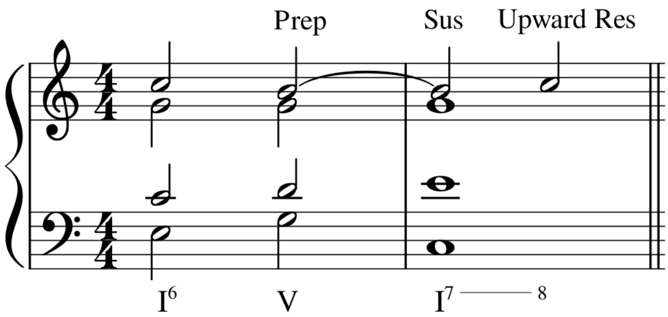
Note. Upper (UN), lower (LN) and double (DN) neighbour tones (from Spyra & Woolhouse, 2021).

Chordal skips are, as the name suggests, chord-tones. These also usually occur on the offbeat, but do not move by step as passing or neighbour tones do (see Figure 3.5 for an example). Instead, they *skip*, meaning they can leap on the staff.

Rhythmic figuration displaces chord-tones in time. These include suspensions, retardations, and anticipations. Suspensions and retardations have three features:

a preparation, suspension, and resolution. Thus, a *prepared* chord-tone (Prep) is *suspended* (Sus) before *resolving* (Res) up or down by a step to a new chord-tone (see Figure 3.2). Suspensions resolve downward and retardations resolve upwards.

Figure 3.2
Retardation in Parts



Note. Example of a retardation, indicating its preparation, suspension (retardation) and upward resolution.

In summary, melodic components of surface features are harmonic in nature and are used to embellish the composition (for a primer, see Aldwell et al., 2010). *Rhythmic components*, on the other hand, add rhythmic activity without disrupting or changing the harmony, i.e., they add busyness without interacting with melodic material. An example of this may be a repeated chord tone, composed in eighth notes, in an otherwise homophonic quarter-note stimulus (see Figure 3.3b).

Perceptually speaking, surface features may strengthen the perception of a non-adjacent key relationship; McAdams (1989) argues that an important part of perception is that deep changes must be reflected in the musical surface as large-scale tonal structures may be perceived weakly without the support of surface features. In fact, listeners seem to be more sensitive to the musical surface than to deeper, harmonic structures (Deliège et al., 1996) and, as a consequence, may use these features to cue memory for deep structures (see also Granot & Jacoby, 2011; Karno

& Konečni, 1992). It is, therefore, important to test the degree to which surface features may influence processing of deeper musical structures.

3.2 Nonadjacent Experiments

Three experiments were conducted to investigate perceptual nonadjacency in the presence of surface features. By systematically varying the features present and by manipulating the duration of the intervening section, these studies tested the degree of influence the musical surface has on completion ratings. Surface features, as described above, consisted of melodic and rhythmic figurations (e.g., passing tones and suspensions) as well as rhythmic activity (the number of notes per harmonic section). I hypothesized that surface features would provide a realistic, and thereby salient, musical experience, facilitating memory for the nonadjacent section.

Experiment 1 tested whether surface features affect ratings of goodness-of-completion compared to baseline. Surface features were broken into *Figuration*, which included both melodic and rhythmic figuration, and *Activity*, both of which could be present or absent from X_{ns}, the nonadjacent section. This led to four main conditions: (1) neither *Figuration* nor *Activity* was present in X_{ns}, (2) only *Figuration* was present, (3) only *Activity* was present, and (4) both *Figuration* and *Activity* were present in X_{ns}. I predicted that the addition of surface features would lead to higher goodness-of-completion ratings. This would result in condition 4 (both are present) acquiring highest ratings, followed by conditions 2 and 3 (only *Figuration* or *Activity* is present), and lastly, condition 1 (no surface features are present). The consistency of surface features between nonadjacent sections (X_{ns}

and X_{pc}) was also tested with the prediction that a consistent musical surface would bolster ratings as well.

Experiment 2 tested harmonic memory by introducing nonadjacent harmonic relationships to the paradigm used in Experiment 1. Again, *Figuration* and *Activity* could be present or absent in X_{ns}. However, in addition to the manipulation of surface features, X_{ns} and X_{pc} could be harmonically congruent or incongruent. This led to eight conditions: each of the four described in Experiment 1 with either harmonically matching nonadjacent sections (X_{ns}-Y_{is}-X_{pc}) or non-matching (Z_{ns}-Y_{is}-X_{pc}). Here, I predicted that the presence of surface features would also increase goodness-of-completion ratings for harmonically congruent conditions more than incongruent conditions. Such a result would suggest that salient musical compositions increase harmonic memory as compared to homophonic stimuli.

Lastly, Experiment 3 investigated the duration of harmonic memory in the presence of surface features. Here, the presence of surface features was held constant, though, once again, the nonadjacent harmonic relationship could be either congruent or incongruent. The intervening section was manipulated in duration such that it could last between 6-36 seconds. Difference scores between congruent and incongruent nonadjacent relationships could provide an estimate of how long the memory for the original, nonadjacent key remains after modulation. I predicted that the presence of surface features in this paradigm would extend the memory decay period beyond the 11 seconds observed in previous research (Woolhouse et al., 2016).

3.3 Experiment 1

Experiment 1 tested whether the addition of surface features influenced goodness-of-completion ratings. Three main factors were used: *Figuration* and rhythmic *Activity* as described above, and *Consistency*, whether surface features were matching between nonadjacent sections of the stimulus. Control factors included *Sequence* (cycle of fifths or non-cycle of fifths chord progressions), modulation *Direction* (either up or down), modulation *Distance* (either 2, 4 or 6 semitones), and nonadjacent *Key* (12 major keys). Factors *Figuration* and *Activity* were chosen for their precise operationalization and the clarity with which one could manipulate both in varying, often integrated, levels. *Consistency* was added to test whether it was necessary for surface features to match between outer sections to create a stable nonadjacent precept. I hypothesized that *Figuration*, *Activity*, and *Consistency* would modify how realistic stimuli would sound and, in turn, influence memory for the nonadjacent section. Specifically, I predicted that they would all lead to higher goodness-of-completion ratings given the conditions heightened familiarity to every-day music.

3.3.1 Methods

Participants

Forty-seven undergraduate university students (18 male, 29 female; ages 17-34, M = 18, SD = 1.7) participated in the experiment. Twenty-one participants self-reported at least 5 years of musical training, including those who received formal lessons and those who were self-taught (M = 9.00; SD = 2.14). A power analysis was conducted for a 2x2x2 repeated measures design using the Superpower package in R (Lakens & Caldwell, 2021). These simulations recommended a minimum of 30 participants. One participant failed to complete the experiment and was excluded from analysis. Participants in this, and all following studies, were given one course credit for participation.

Apparatus and Procedure

Stimuli were composed using the open-source software *MuseScore2* (MuseScore Project, 2015) and saved as MIDI files. The software's synthesized bassoon timbre was used for its quick attack rate and sustained amplitude envelope. This balance was necessary for stimuli that included suspended notes as the perception of these would have been lost in a timbre with a fast decay, e.g., piano. Ninety-six stimuli were created with a systematic combination of each condition within *Figuration*, *Activity*, *Consistency*, *Sequence*, *Direction*, and *Distance* (i.e., 2x2x2x2x2x3 levels of each factor).

Participants gave informed consent prior to beginning the experiment and completed a demographic questionnaire. A program was created for stimulus presentation using Python 3.6 and the *Kivy 1.9* GUI package (Kivy Organization, 2016). Participants used AKG K 172 HD headphones (frequency range 18 Hz–26 kHz) for the duration of the experiment. Two novel practice stimuli were presented to ensure participants understood the task. In the main experiment, stimuli were presented in a unique random order to each participant, with a randomized transposition per stimulus. Participants were asked to rate the probe cadence on its goodness-of-completion using a 7-point Likert-type sliding scale presented by the Python program, labelled on both ends (1 = *not at all* and 7 = *strong sense of* *completion*). A continuous measurement was chosen to encourage the participant to interpret the scale as interval and not ordinal (Howell, 2016, pp. 21-24).

Stimuli

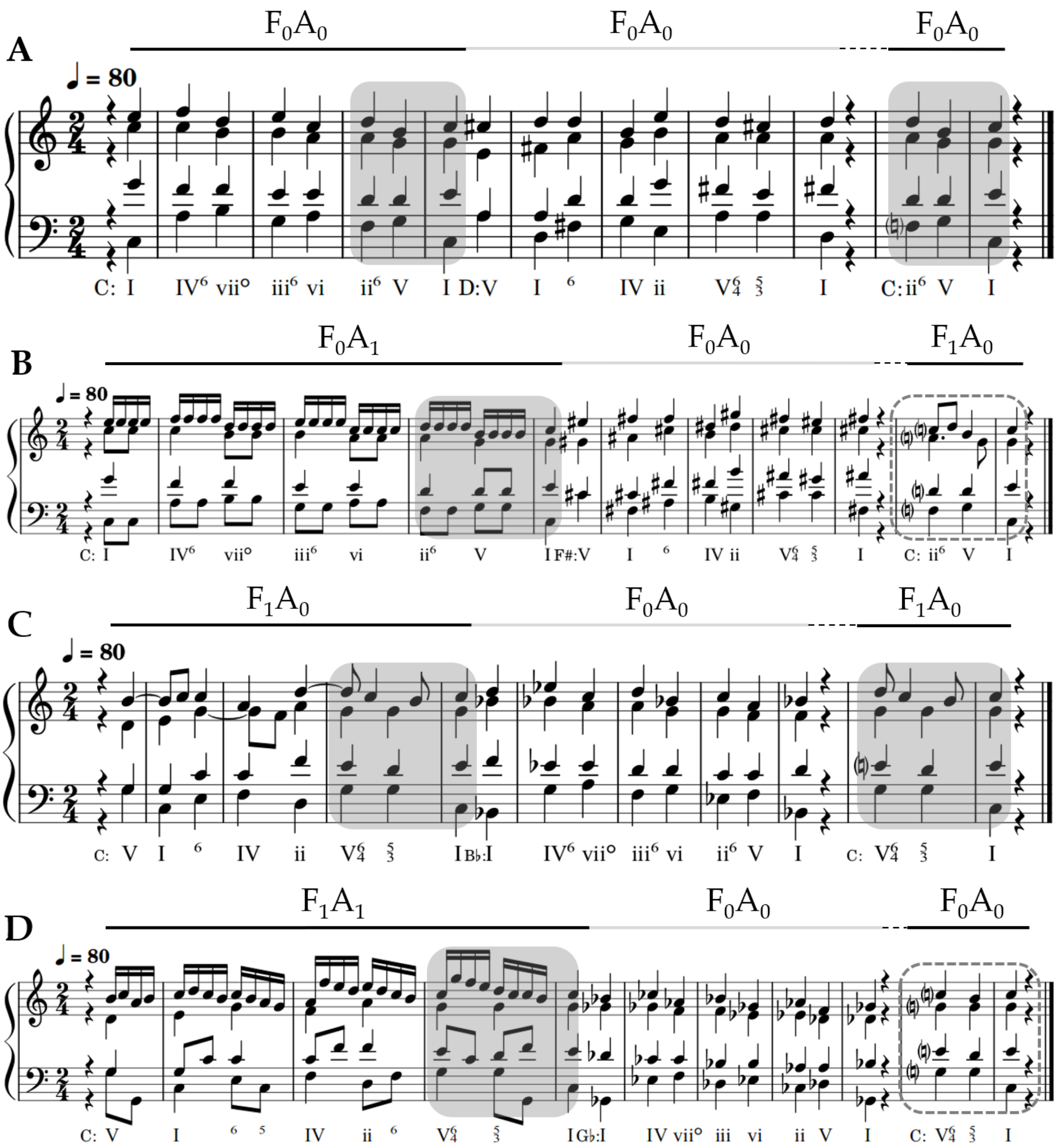
Stimuli were composed using the nonadjacent paradigm discussed at length in Chapter 2. Nonadjacent and intervening sections were each composed of eight chords (lasting 6s), after which there was a single-beat rest (0.75s) followed by the probe cadence (see Figure 3.3). Three chords were used in the probe cadence to ensure participants recognized the key adequately (see Spyra et al., 2021 for an overview of local effects). In full, stimuli lasted for 15s each (20 beats at 80 BPM). Nonadjacent sections (X_{ns} and X_{pc}) were composed in the same key. All 12 major keys of Western tonal-harmonic music were used. The intervening section was modulated up or down two, four, or six semitones.

Factors

Figuration and *Activity* were either present (F_1, A_1) or absent (F_0, A_0) in the non-adjacent section (X_{ns}). When *Figuration* was present, the section was composed with the inclusion of either melodic or rhythmic figurations. Conversely, when it was absent, X_{ns} was homophonic. When *Activity* was present, chord-tones were added as repeated eighth and sixteenth notes. In summary, the simplest version mirrored stimuli in Woolhouse et al. (2016) or Spyra et al. (2021) and were composed of homophonic quarter-note chords (F_0A_0; Figures 3.3a, 3.4a). However, if the condition had *Activity* present, but not *Figuration* (F_0A_1; Figures 3.3b, 3.4b), the stimulus was composed with repeated sixteenth notes and eighth notes.

Figure 3.3

Stimuli in all Factor Combinations



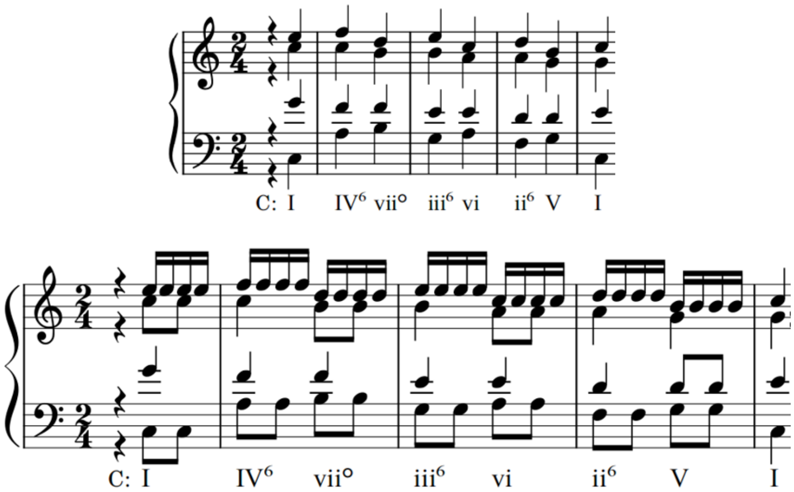
Note. Figuration (F) and *Activity* (A) are marked with subscripts denoting the presence (1) or absence (0) of each factor. Highlighted and dashed regions indicate consistent or inconsistent surface features between nonadjacent sections. From Spyra and Woolhouse (2021).

Figure 3.4

Two Conditions: F_0A_0 and F_0A_1

A

B



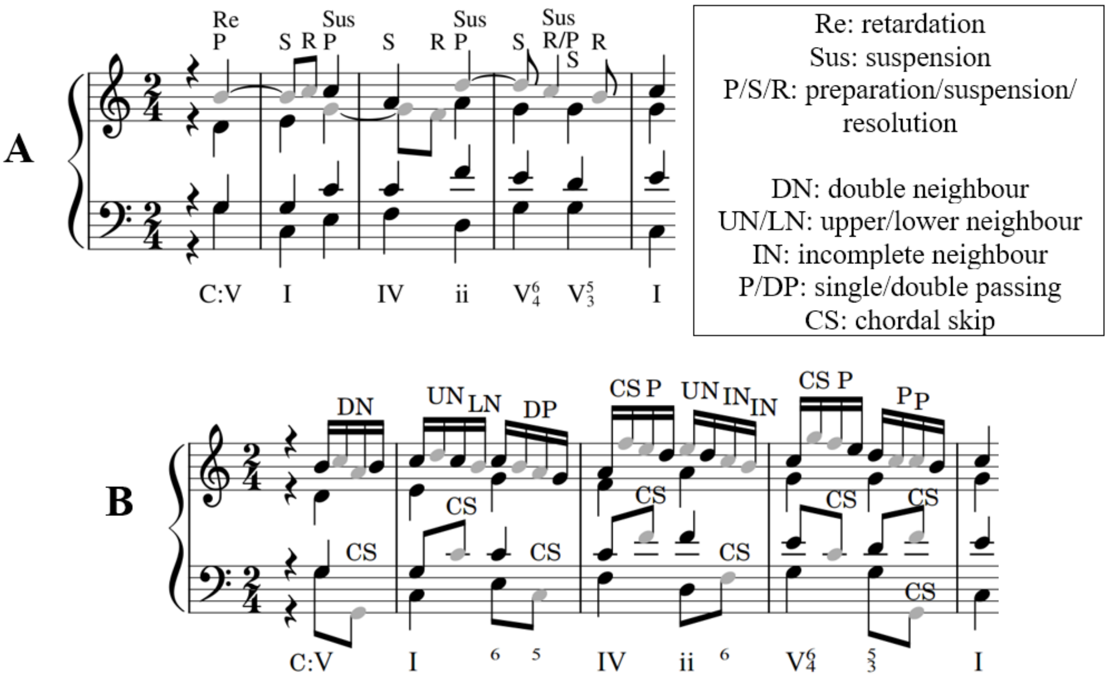
Note. Sample stimuli with (a) *Figuration* and *Activity* absent (i.e., the F_0A_0 condition) and (b) with *Figuration* absent, but *Activity* present (i.e., F_0A_1 condition). From Spyra and Woolhouse (2021).

Stimuli that included *Figuration* but not *Activity* (F_1A_0) contained suspensions and retardations, providing melodic interest without increasing the number of notes (Figures 3.3c, 3.5a). Lastly, stimuli with both *Figuration* and *Activity* (F_1A_1) were composed using sixteenth and eighth notes like in F_0A_1, but also employed passing tones, chordal skips, and neighbour tones in a melodic arrangement (Figures 3.3d, 3.5b). The intervening section, Y_{is}, was always homophonic (F_0A_0).

Lastly, the factor, *Consistency*, pertained to the similarity of surface features between the nonadjacent section and the probe cadence (i.e., X_{ns} and X_{pc}). These could be either the same (PC_{same}) or different (PC_{diff}). When *Consistency* was different, *Figuration* and *Activity* were complete reversals of surface features between the nonadjacent sections. For example, if the nonadjacent section was composed with both *Figuration* and *Activity* present (F_1A_1), and *Consistency* was different,

Figure 3.5

Two Conditions: F_1A_0 and F_1A_1



Note. Sample stimuli with (a) *Figuration*, but not *Activity* absent (i.e., the F_1A_0 condition) and (b) with both *Figuration* and *Activity* (i.e., F_1A_1 condition). From Spyra and Woolhouse (2021).

surface features would both be absent (F_0A_0) in the probe cadence. Similarly, if *Figuration* was present but *Activity* was not (F_1A_0), the probe cadence would have the opposite configuration: absent *Figuration*, but present *Activity* (F_0A_1).

Three additional factors (*Sequence*, *Direction* and *Distance*) were added to increase generalizability, to combat familiarity and, consequently, participant fatigue. The harmonic progression itself (i.e., *Sequence*, the order of chords chosen for the sequence) varied between stimuli; each section was composed in a cycle of fifths (S_{c5}: I-IV-vii°-iii-vi-ii-V-I) or a regular, well-formed harmonic progression (S_{reg}: V-I-I^6-IV-ii-V$^6_4$-V$^5_3$-I). The nonadjacent and intervening sections always contrasted each other in this regard: if X_{ns} was composed using a S_{c5} progression, Y_{is}

was composed in an S_{reg} progression and vice versa. The probe cadence always matched the last three harmonies of X_{ns} (S_{c5}: ii-V-I; S_{reg}: V^6_4-V^5_3-I). Lastly, modulations between X_{ns} and Y_{is} moved up or down (*Direction*) by 2, 4, or 6 semitones (*Distance*).

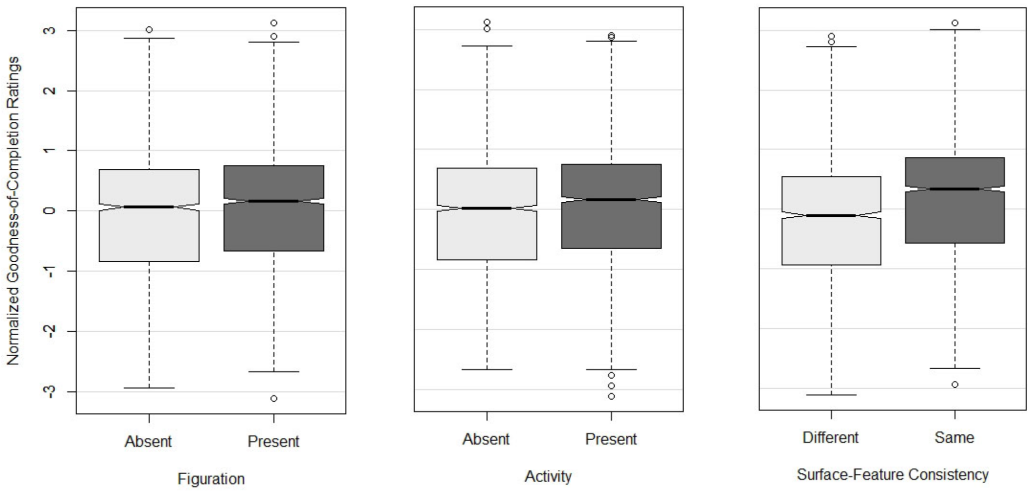
3.3.2 Results

A 2x2x2 repeated measures ANOVA was conducted using raw scores with factors *Figuration* (F_0, F_1), *Activity* (A_0, A_1), and *Consistency* (PC_{same}, PC_{diff}). All three main effects reached significance: *Figuration* ($F_{1,45} = 8.75$, $p = 0.005$, $\eta^2_p = 0.163$), where F_1 was rated higher than F_0 (M: $F_1 = 0.048$; $F_0 = -0.048$); *Activity* ($F_{1,45} = 22.33$, $p < 0.001$, $\eta^2_p = 0.332$) where, similarly, A_1 was rated higher than A_0 (M: $A_1 = 0.063$; $A_0 = -0.063$); and *Consistency* ($F_{1,45} = 41.03$, $p < 0.001$, $\eta^2_p = 0.477$), where PC_{same} was rated higher than PC_{diff} (M: $PC_{same} = 0.161$; $PC_{diff} = -0.161$; see Figure 3.6).[1]

The interaction between *Activity* and *Consistency* was also significant ($F_{1,45} = 7.10$, $p < 0.05$, $\eta^2_p = 0.136$); participants rated PC_{same} consistently higher than PC_{diff}, whether *Activity* was present or absent, but rated PC_{diff} higher when *Activity* was present than when it was absent (PC_{same}: $M_{A0} = 0.172$, $M_{A1} = 0.151$; PC_{diff}: $M_{A0} = -0.299$, $M_{A1} = -0.024$; Figure 3.7). This suggests that there may be a memory boost for stimuli that begin with busier sequences than the homophonic progressions found in previous literature.

[1] Normalized scores were used for visual purposes only with the purpose of balancing all rating styles. Raw data were used for all analyses.

Figure 3.6
Three Main Effects: Figuration, Activity, and Consistency



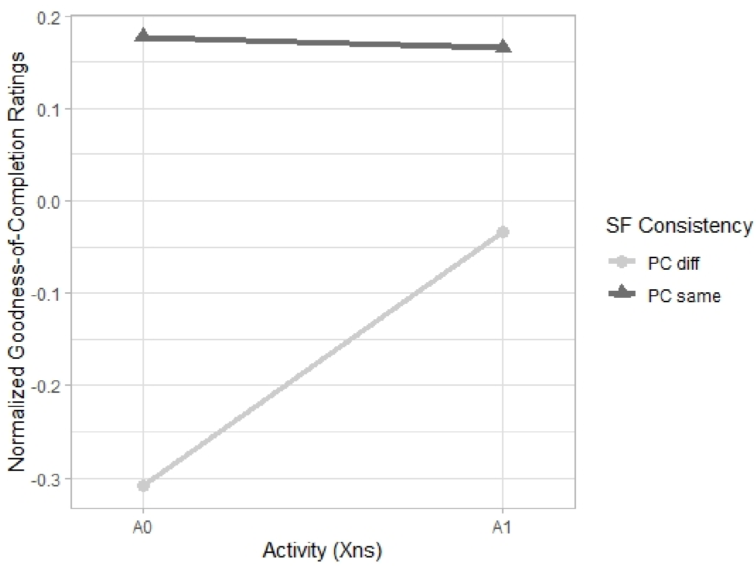
Note. Normalized goodness-of-completion ratings for *Figuration*, *Activity* and *Consistency* showing medians and 95% confidence intervals. From Spyra and Woolhouse (2021).

The interaction between *Activity* and *Figuration* did not reach significance ($p = 0.46$), nor did the interaction between *Figuration* and *Consistency* ($p = 0.68$). Lastly, the 3-way interaction between *Figuration*, *Activity*, and *Consistency* also did not reach significance ($p = 0.55$).

To summarize, the homophonic condition (F_0A_0) elicited lowest responses ($M = -0.089$), followed by F_1A_0 ($M = -0.038$), F_0A_1 ($M = -0.008$), and, lastly, F_1A_1 ($M = 0.135$).The difference between ratings for PC_{same} and PC_{diff} was over twice as large as that for levels in either *Figuration* or *Activity* (M: $PC_{same} = 0.161$; $PC_{diff} = -0.161$; difference score = 0.322; *Figuration* difference = 0.096; *Activity* difference = 0.126).

Musical experience was compared in post-hoc analysis. The between-groups factor, *Musician*, was added to the repeated-measures ANOVA described above.

Figure 3.7
Interaction between Activity and Consistency



Note. Mean normalized goodness-of-completion (GoC) ratings with standard error. From Spyra and Woolhouse (2021).

No interaction between it and any other independent variable reached significance, suggesting that (self reported) musical experience does not influence participants' ability to perceive nonadjacent surface feature relationships.

3.3.3 Discussion

Experiment 1 examined whether surface features could augment the perception of structural coherence in music. It was hypothesized that including surface features and keeping them consistent between X_{ns} and X_{pc} would affect goodness-of-completion ratings due to heightened salience and generalizability of the stimuli. Results supported this hypothesis: ratings were higher for stimuli that included *Figuration* and/or *Activity*, and more so for stimuli where surface features were consistent between nonadjacent sections. This suggests that higher degrees of generalizability in the stimuli, as well as stylistic consistency between target and cue sections, is crucial for the perception of musical form.

In the interaction between *Activity* and *Consistency*, *Activity* had no significant effect on ratings as long as surface features were consistent between nonadjacent sections. However, when surface features were not consistent, the absence of *Activity* in the nonadjacent section significantly reduced ratings. These results suggest that (1) the replication of the musical surface is vital for the perception of nonadjacency and that, (2) without consistency, a higher degree of activity increases memorability for a musical passage, perhaps due to a greater familiarity for music with higher rhythmic activity. It is possible that these results are reflective of a stronger representation of schemata in long-term memory. For those participants to whom music with surface features is more familiar, such stimuli would fit into schemata built through enculturation and may, therefore, be easier to process (see Agres, 2019; Bey & McAdams, 2002).

3.4 Experiment 2

Experiment 2 tested the effects of *Figuration* and *Activity* on harmonic coherence and memory by juxtaposing congruent and incongruent nonadjacent key relationships, similar to methods used in Woolhouse et al. (2016). Through this juxtaposition, one can extract local and global harmonic effects and, thereby, highlight the differences between harmonic memory in homophonic conditions compared to those with surface features present. I hypothesized that increased surface features would create a more generalizable musical experience and would, therefore, augment memory for key. I further predicted that, if this hypothesis was supported by data, surface features would illicit higher goodness-of-completion ratings for congruent conditions than for incongruent conditions.

3.4.1 Methods

Participants

Eighty-three undergraduate university students participated in this experiment (68 female, 33 male, 6 no response; ages 17-32, $M = 19$, $SD = 2.3$). Twenty-six self-identified as musicians with 5+ years of musical training ($M = 7.6$, $SD = 2$). Once again, one participant failed to complete the experiment and was excluded from analysis.

Apparatus and Procedure

Stimuli were composed in *MuseScore3* (MuseScore Project, 2015) and exported as wav files. The software's synthesized flute timbre was used as it was similar to the bassoon in Experiment 1 in terms of attack rate and sustained amplitude envelope but was deemed more pleasant. Ninety-six stimuli were created, in all combinations of six factors: *Figuration*, *Activity*, *Consistency*, *Nonadjacent Key Relationship*, *Direction*, and *Distance* (i.e., 2x2x2x2x2x3 levels). The experiment was built in *PsychoPy3* and *PsychoJS* (Peirce et al., 2019), and was hosted on *Pavlovia* (https://pavlovia.org/). Due to the Covid-19 pandemic, Experiment 2 was run virtually. Participants, therefore, used personal devices and headphones. Participants completed the demographic survey online, using *LimeSurvey* (Limesurvey

GmbH./LimeSurvey, 2020), which included the consent form. All other aspects of the procedure were the same as in Experiment 1.

Stimuli

Stimuli used in Experiment 2 were similar to those in Experiment 1: factors included *Figuration* (present, absent), *Activity* (present, absent), and *Consistency* (same, different). However, an additional factor, *Nonadjacent Key Relationship* (congruent, incongruent), was included to test harmonic memory in the presence of surface features. A stimulus with a congruent *Nonadjacent Key Relationship* began and ended in the same key (i.e., X_{ns}-Y_{is}-X_{pc}). However, a stimulus with an incongruent *Nonadjacent Key Relationship* had no harmonic completion and was composed in three different keys (i.e., Z_{ns}-Y_{is}-X_{pc}, see Figure 3.8).[2] Modulations were labeled backwards from Z to X for ease of comparison between conditions as X_{pc} was the target section of all stimuli. All additional factors remained the same.

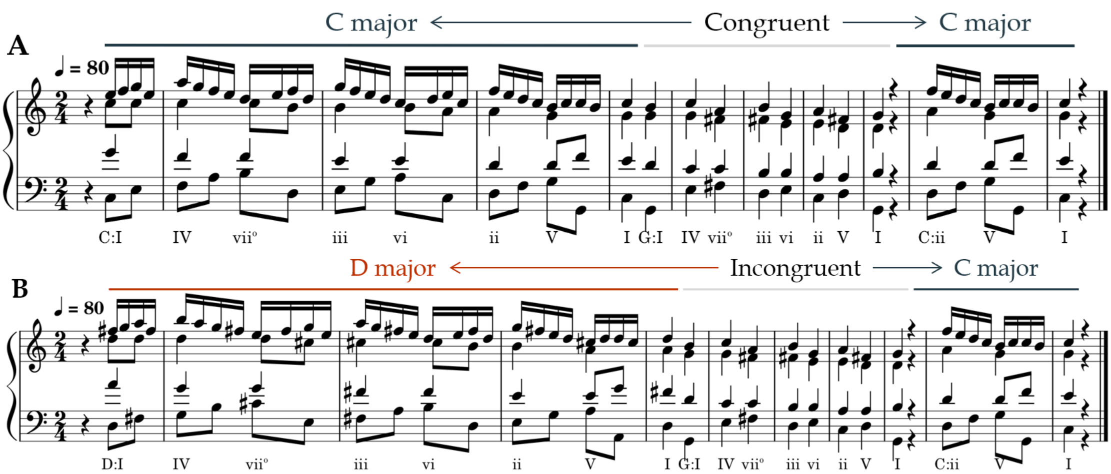
3.4.2 Results

A 2x2x2x2 repeated-measures ANOVA was used to analyze four factors: *Figuration* (present, absent), *Activity* (present, absent), *Consistency* (same, different), and *Nonadjacent Key Relationship* (congruent, incongruent). Once again, the main effect of *Figuration* was significant ($F_{1,81} = 7.57$, $p < 0.05$, $\eta_p^2 = 0.085$); F_1 was rated higher than F_0 (M: $F_1 = 0.03$; $F_0 = -0.03$). *Nonadjacent Key Relationship* was also significant ($F_{1,81} = 30.21$, $p < 0.001$, $\eta_p^2 = 0.272$); congruent relationships were rated higher than incongruent (M: congruent = 0.07; incongruent = -0.07;

[2] Once again, these letters relate to the harmonic relationship in relation to the target section of the stimulus, the probe cadence. Hence, Z-Y-X signifies that Y is some distance from X and, consequently, Z is another such distance from Y.

Figure 3.8

Stimuli in Both Levels of Factor: Nonadjacent Key Relationships



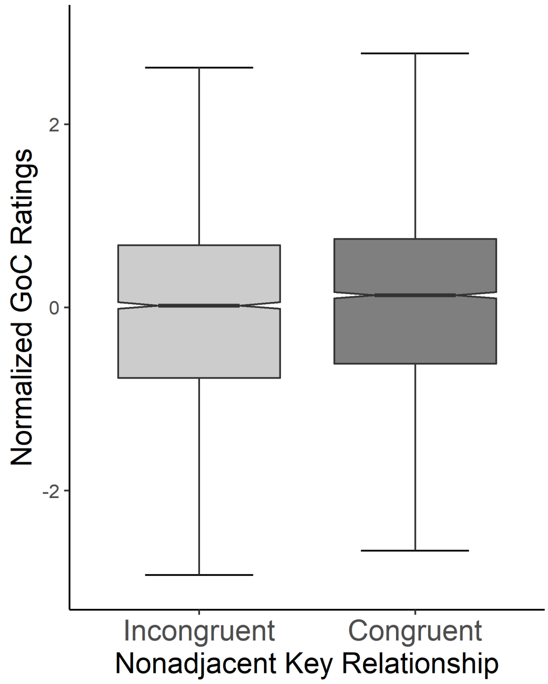
Note. Juxtaposition of *Nonadjacent Key Relationships* where (A) is congruent, i.e., X_{ns}-Y_{is}-X_{pc}, and (B) is incongruent, i.e., Z_{ns}-Y_{is}-X_{pc}.

Figure 3.9). However, main effects of *Activity* (p = 0.80) and *Consistency* (p = 0.23) did not reach significance.

There was a significant spreading interaction between *Figuration* and *Activity* ($F_{1,81} = 11.85$, $p < 0.001$, $\eta_p^2 = 0.128$) in which conditions where both factors were present (F_1A_1; Figure 3.10a) showed the highest ratings (F_0: $M_{A0} = 0.031$, M_{A1} = -0.089; F_1: M_{A0} = -0.022, $M_{A1} = 0.079$). The interaction between *Figuration* and *Consistency* was also significant ($F_{1,81} = 9.54$, $p < 0.01$, $\eta_p^2 = 0.105$); ratings were highest for conditions where *Figuration* was present, and *Consistency* was the same (PC_{same}: M_{F0} = -0.079, $M_{F1} = 0.108$; PC_{diff}: $M_{F0} = 0.022$, M_{F1} = -0.050; Figure 3.10b). Lastly, the crossover interaction between *Activity* and *Consistency* was also significant ($F_{1,81} = 8.53$, $p < 0.01$, $\eta_p^2 = 0.095$); ratings were, again, highest when *Activity* was present, and *Consistency* was the same (PC_{same}: M_{A0} = -0.029, $M_{A1} = 0.058$; PC_{diff}: $M_{A0} = 0.039$, M_{A1} = -0.067; Figure 3.10c).

Figure 3.9

Main Effect: Nonadjacent Key Relationship

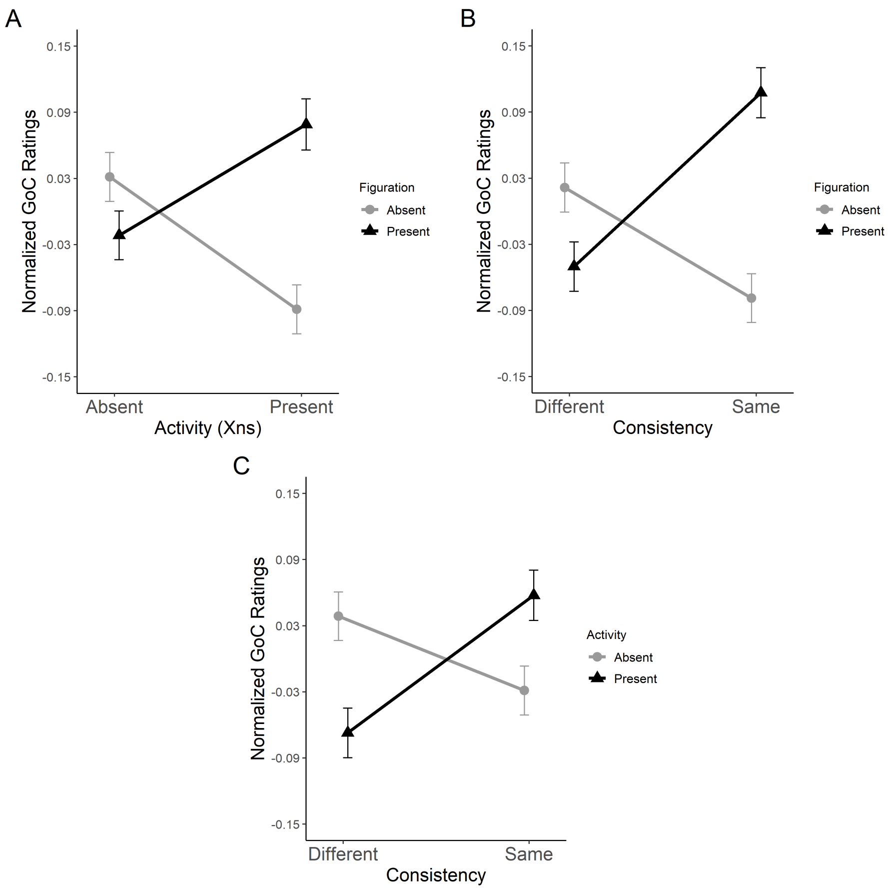


Note. Goodness-of-completion (GoC) ratings between congruent and incongruent key relationships using medians and 95% confidence intervals (from Spyra & Woolhouse, 2021, p. 20).

The three-way interaction between *Nonadjacent Key Relationship*, *Figuration*, and *Consistency* was significant ($F_{1,81} = 4.52$, $p < 0.05$, $\eta_p^2 = 0.053$). When the *Nonadjacent Key Relationship* was congruent, goodness-of-completion ratings were higher, on average (Figure 3.11a), than incongruent relationships (Figure 3.11b), demonstrating the factor's significant main effect as discussed above. Furthermore, when *Nonadjacent Key Relationship* was congruent (Figure 3.11a), the presence of consistent figuration (i.e., condition F_1, PC_{same}) had a significantly higher mean rating ($M_{F1,PCsame} = 0.173$) than other conditions ($M_{F0,PCsame} = 0.003$, $M_{F0,PCdiff} = 0.055$, $M_{F1,PCdiff} = 0.035$). This is in line with my hypothesis as there seems to be a significant memory boost when *Figuration* is present and consistent within the

Figure 3.10

2-way Interactions: Activity, Figuration, and Consistency

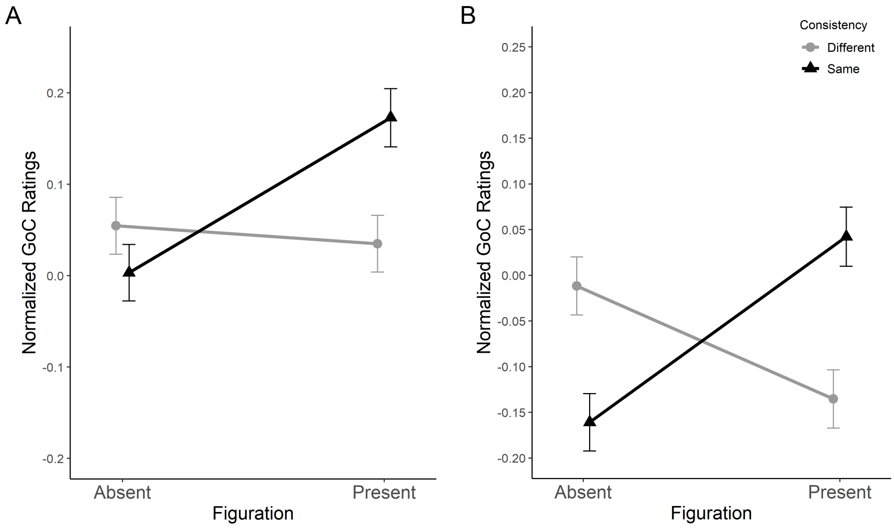


Note. From Spyra and Woolhouse (2021, p. 21).

congruent key relationship condition where memory should be highest (i.e., X_{ns}-Y_{is}-X_{pc} where nonadjacent sections match). In fact, congruent key relationships should elicit consistent results as they have a clear perceptual beginning, middle, and end (i.e., none of the points in Figure 3.11a should be significantly different). However, the fact that the addition of figuration boosts ratings of completion provides strong evidence that the surface features, especially when consistent between nonadjacent sections, play a vital role in harmonic memory. In other words, memory for tonal structures increases in the presence of figuration.

Figure 3.11

Three-way Interaction: Figuration, Consistency, and Nonadjacent Key Relationships



Note. Means and standard error of goodness-of-completion (GoC) ratings of (a) congruent key relationships and (b) incongruent key relationships (from Spyra & Woolhouse, 2021, p. 23).

The interactions between *Nonadjacent Key Relationship* and *Figuration* (p = 0.50), *Nonadjacent Key Relationship* and *Activity* (p = 0.21), and between *Nonadjacent Key Relationship* and *Consistency* (p = 0.97) did not reach significance. Similarly, 3-way interactions between *Nonadjacent Key Relationship*, *Figuration*, and *Activity* (p = 0.70); *Nonadjacent Key Relationship*, *Activity*, and *Consistency* (p = 0.18); and between *Figuration*, *Activity*, and *Consistency* (p = 0.97) did not reach significance. Lastly, the 4-way interaction between *Nonadjacent Key Relationship*, *Figuration*, *Activity*, and *Consistency* also did not reach significance (p = 1.00).

As in Experiment 1, *Musician* did not significantly interact with any other independent variable in post-hoc analysis.

3.4.3 Discussion

Participants, regardless of musical experience, were able to distinguish between congruent and incongruent nonadjacent tonal relationships suggesting they could remember musical keys despite the intervening harmonic sequences. Though the main effects for both *Activity* and *Consistency* did not reach significance—a surprising result given they were so prominent in Experiment 1—the unpredictability of collecting data virtually, particularly during a global pandemic, has likely introduced noise into the results. A replication of this experiment should be run in person to verify the current findings.

The significant two-way interaction between *Figuration* and *Activity* provides evidence suggesting that surface features increase goodness-of-completion. Further, the interactions between *Consistency* and *Figuration* or *Activity* suggest that the consistency of surface features between nonadjacent sections increases the sense of structural completion. Finally, the three-way interaction between *Figuration*, *Consistency* and *Nonadjacent Key Relationship* suggests that perception and memory of harmonic structures is augmented by the presence of surface features; congruent key relationships were rated highest when figuration was present and consistent. These results support and extend the findings of Experiment 1 and the research discussed in Chapter 2.

3.5 Experiment 3

In Experiment 3, the duration of the intervening section was manipulated to last well beyond that found in Farbood (2016) and Woolhouse et al. (2016). Additionally, key relationships were manipulated—as in Experiment 2—to calculate the duration at which the difference between congruent key relationships and incongruent ones decayed to zero. I predicted that surface features will bolster harmonic memory, i.e., differences between stimuli with congruent and incongruent key relationships will still be found past the 11 second time limit determined in Woolhouse et al. (2016).[3] An 11-second target was chosen as both Experiment 3 and Woolhouse et al. (2016) employ diatonic intervening sections, whereas results from Farbood (2016) rely on single-chord intervening sections that may not be comparable.

3.5.1 Methods

Participants and Procedure

Ninety-three participants were recruited from the universities undergraduate research pool (65 female, 28 male) with ages ranging from 18-65 ($M = 20$, $SD = 5.6$). The same procedure was followed as in Experiment 1, with the exception of stimulus creation, outlined below.

Stimuli were generated in *MuseScore2* (MuseScore Project, 2015) and exported as MP3 files. The default grand piano synthesizer was used for this experiment as

[3]For clarity, I was predicting a significant spreading interaction between *Nonadjacent Key Relationship* and *Duration* where congruent and incongruent key relationships would begin significantly different at lower durations and the difference between them would steadily decrease as duration increased.

the stimuli did not require special timbral properties to aid perception. Thirty-six stimuli were presented to each participant in a randomized order resulting in a 20-minute study in total. See Experiment 1 for apparatus and programming details.

Stimuli

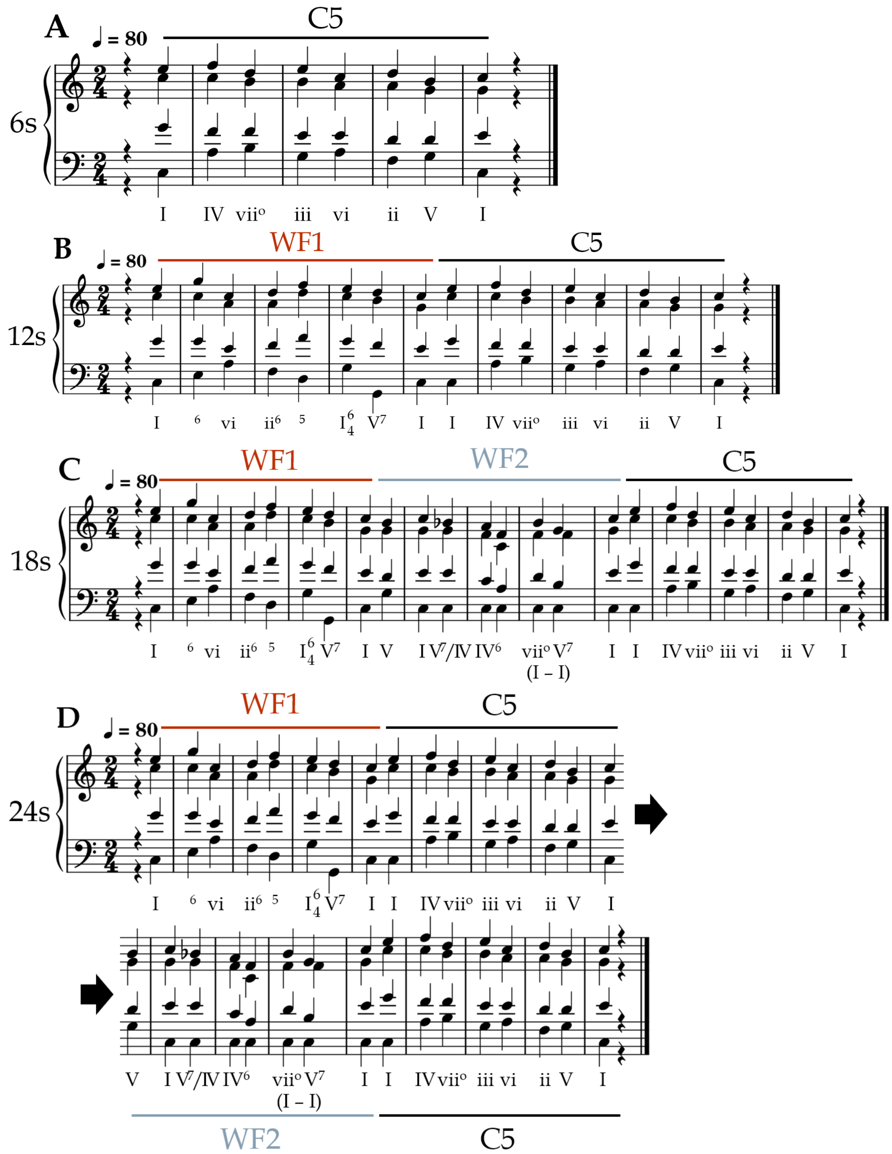
Stimuli resembled the F_1A_1 condition in Experiments 1 and 2; nonadjacent sections were exact replicas of conditions with both *Figuration* and *Activity* present. The intervening section was always homophonic in nature and was modified to last 6, 12, 18, 24, 30, and 36 seconds in duration (Figure 3.12). The 6-second sequence was composed using a cycle-of-fifths progression. Subsequent additions alternated between cycle-of-fifths and well-formed progressions, all following standard practices of Western tonal music (see Figure 3.12; Piston, 1941). In addition to the factor, *Duration* (6-36 sec), the *Nonadjacent Key Relationship* was again modified to be either congruent (X_{ns}-Y_{is}-X_{pc}) or incongruent (Z_{ns}-Y_{is}-X_{pc}).

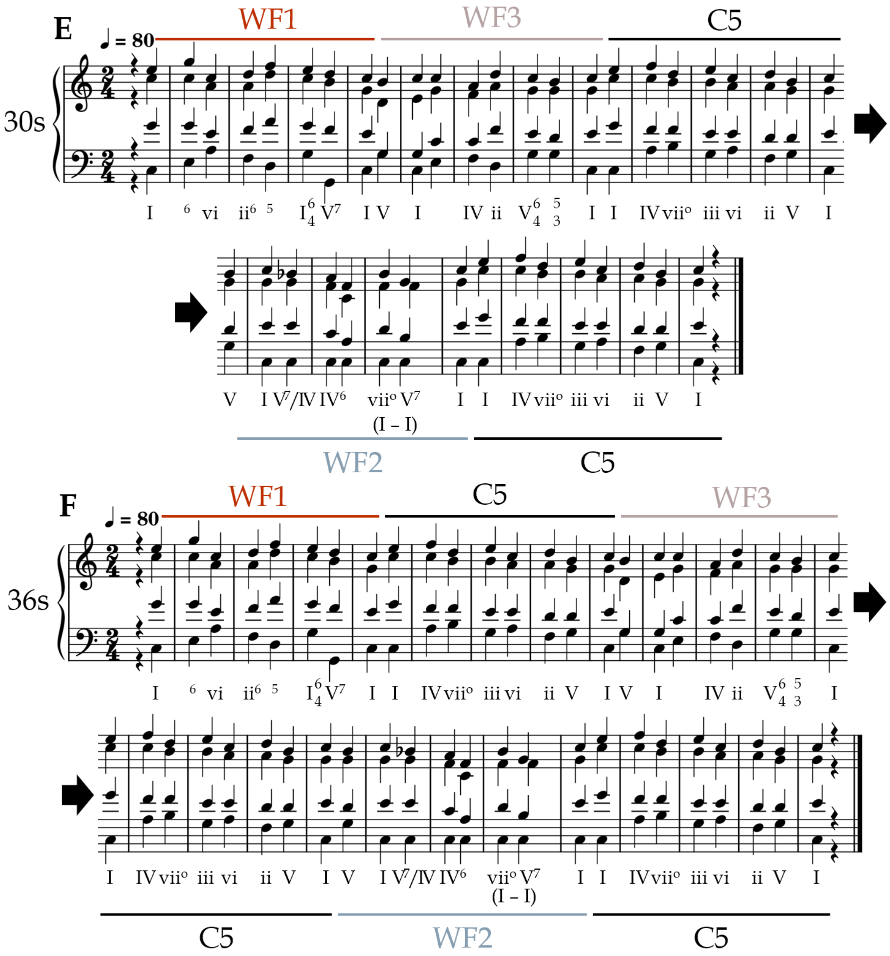
All additional factors were similar to those in Experiments 1 and 2: keys were randomly assigned to probe cadences. All 12 major keys found in Western tonal-harmonic music were used. Each stimulus section modulated up or down either one, two, or three steps on the cycle-of-fifths using the same methods as in Experiment 2[4].

[4]The probe cadence (X_{pc}) was considered as the target key of the stimulus. Thus, modulations were calculated backwards as in Z_{ns}-Y_{is}-X_{pc}

Figure 3.12

Intervening Sections by Duration





Note. A new subsection was added to each duration, using both *well-formed* (WF) progressions and *cycle-of-fifths* (C5) progressions.

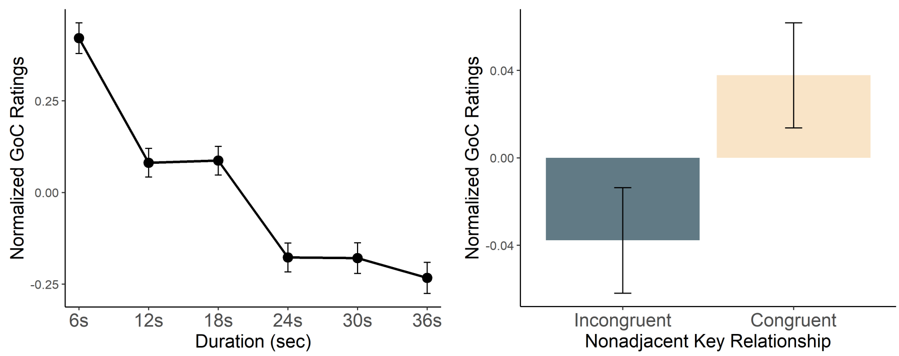
3.5.2 Results

A linear mixed effects modeling approach was used due to the complexity of the data. This approach allows for repeated measures designs which violate the assumption of independence. It also allows the model to have a mixture of fixed effects (i.e., variables with repeatable and fixed levels) and random effects (i.e., those randomly sampled from a population). Lastly, it allows the analyst to have independent variables which are continuous (e.g., *Duration*) and ones that are categorical (e.g., *Nonadjacent Key Relationship*) in the same analysis.

Using *lme4*, a dedicated package for R (Bates et al., 2014), the organization of a typical ANOVA was followed to fit separate models for each main effect and interaction. *Duration* and *Nonadjacent Key Relationship* were included as fixed effects in models 2 and 3 respectively and their interaction was added in model 4. Fixed effects are interpreted like main effects in an ANOVA. *Participants* were included in all models as random effects with random intercepts—uncontrolled by the experimental design—to account for between-participant variance in rating style. More information on the models and details of the results can be found in Tables A1.1, A1.2, and A1.3 of Appendix A.

There was a significant main effect of *Duration* (b = -0.03, t = -0.03, $p < 0.001$) where ratings decreased as the duration increased (M_{6s} = 0.421, M_{12s} = 0.081, M_{18s} = 0.087, M_{24s} = -0.177, M_{30s} = -0.179, M_{36s} = -0.233; Figure 3.12). There was also a significant main effect of *Nonadjacent Key Relationship* (b = 0.11, t = 2.2, $p < 0.05$); congruent harmonic relationships were rated higher, on average, than incongruent relationships ($M_{congruent}$ = 0.038, $M_{incongruent}$ = -0.038; Figure 3.12).

Figure 3.12
Main Effects: Duration and Nonadjacent Key Relationship



Note. Normalized goodness-of-completion (GoC) ratings using means and standard error bars.

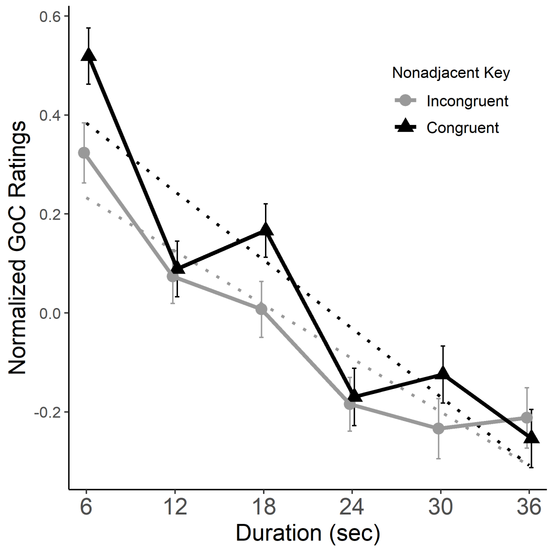
Model 4, which introduces the interaction, approached but did not significantly increase the fit of the model and was dropped from the analysis (M4: $p = 0.09$). A visual inspection of the interaction shows a trend towards a spreading interaction, tentatively supporting the original hypothesis (Figure 3.13): participants are able to tell apart congruent from incongruent stimuli only up to a certain duration, after which memory for the nonadjacent key has decayed. Making the model more complex (by adding the interaction) did not significantly improve its fit. However, Akaike's Information Criterion (AIC)[5] was lowest for Model 4 and the Bayesian Information Criterion (BIC) was lowest for Model 2 making the selection of the model less clear cut. Prior to analysis, I predicted that there would be a significant interaction between *Nonadjacent Key Relationship* and *Duration*. Model 4, with its low AIC, supports this prediction. More studies should, however, be run before making a confident decision on which model best represents the data. Adding *Musician* similarly did not significantly increase the fit of the model.

Difference scores were calculated between congruent and incongruent nonadjacent key relationships for each duration in a post hoc analysis (as per Woolhouse et al., 2016). Previous research found that difference scores decreased as intervening time increased. By fitting a model to these scores, researchers can estimate the limits of how long participants could maintain a harmonic structure in memory. Once the trendline crosses zero, participants could no longer tell the difference between congruent and incongruent key relationships and, presumably, have no memory of the original key left.

[5]AIC provides an estimate of how much information will be lost if a model is used instead of true data. This is a relative number used to compare models, and not a measure of model strength. It also takes into account under- and overfitting (for more information, see Fiquet et al., 2021).

Figure 3.13

Interaction between Nonadjacent Key Relationship and Duration



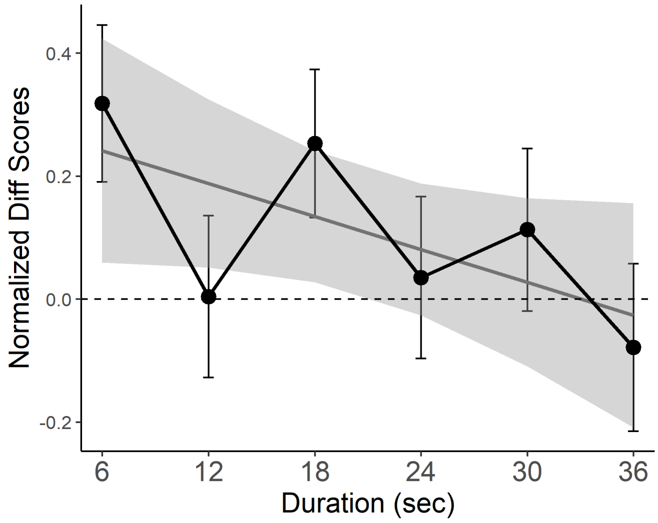
Note. Normalized goodness-of-completion (GoC) scores with standard error bars and the line of best fit.

The equation of a fitted linear model was solved for *Duration* to determine where the difference crossed zero (Figure 3.14). The resulting model crossed the x-axis at 33.04 s. This suggests that participants were able to keep the memory of the nonadjacent key active up to approximately 33 seconds, well past the intercept found by Woolhouse et al. (2016).

3.5.3 Discussion

This study explored memory decay in the presence of surface features. Results support the prediction that including surface features extends the duration of memory for harmonic structures. The factor, *Duration*, showed that there was a decay in memory over time; as duration spent in the intervening key increased, completion scores decreased. *Nonadjacent Key Relationship* provided evidence

Figure 3.14
Difference Scores across Duration



Note. Difference scores between congruent and incongruent conditions with standard error bars, 95% confidence intervals in grey, and a generalized linear model (x-intercept = 33.04).

suggesting participants could reliably distinguish between congruent and incongruent key relationships. Furthermore, congruent relationships were rated higher than incongruent ones, providing support for the assumption that there is, indeed, memory for harmonic structures. The interaction between these factors was only approaching significance but difference scores between congruent and incongruent key relationships per level of *Duration* showed a similar trend as the main effect of *Duration* itself: as the duration of the intervening section increased, difference scores decreased and eventually crossed zero at approximately 33 seconds. This is an interesting finding as difference scores in Woolhouse et al. (2016) crossed zero at approximately 11 seconds. Diatonic conditions in Farbood (2016) also failed to produce differences after 10 seconds whereas single-chord conditions reached 20 seconds, both well below the threshold found in the current study.

This study supports Spyra et al. (2021), who also found that harmonic memory decays over time. Such results are contrary to those found in Lewandowsky et al. (2004), for example, which argue against time-based forgetting. It is possible that the tasks in Lewandowsky et al. (2004) and Experiment 3 targeted different aspects of memory as one employed serial recall and the other a nonadjacent paradigm (see Cowan & Aubuchon, 2008). More studies must be done, however, to distinguish the nuances of this difference.

Such contrasting findings identify a limitation of the current study; goodness-of-completion is only one measure of musical memory. Other measures, such as Farbood's (2016) tension ratings could be used to further support these results as they, like nonadjacent key relationships, depend on the participant sensing the return to a key. It would be interesting to see whether nonadjacent key paradigms would be affected by employing a more direct measure of harmonic memory such as asking participants to reproduce the nonadjacent key (which is, admittedly, difficult, particularly for nonmusicians).

Lastly, the findings of Experiment 3 support Experiments 1 and 2, all of which suggest that more salient stimuli, i.e., those containing surface features, strengthen the resilience of memory for musical keys. This is an important steppingstone in the study of large-scale tonal structures as it can inform future stimulus creation and the directions taken by memory researchers employing the nonadjacent paradigm.

3.6 General Discussion

The experiments described in this chapter make it apparent that surface features and textural consistency between nonadjacent sections are integral to our perception of large-scale tonal structures. In Experiment 1, we saw evidence supporting the idea that surface features significantly change our sense of musical completion. Experiment 2 expanded on this finding by testing memory for harmonic structures and, there too, found significant increases. Finally, Experiment 3 tested the duration of these increases and found that the addition of surface features—when compared to the homophonic stimuli so often used in music-cognition experiments—extended memory decay well beyond that found in previous literature.

As mentioned in the beginning of this chapter, such findings may be considered counterintuitive. Music analysists commonly strip away surface features until only the background structure or *Ursatz* remains (Bharucha, 1994; Lerdahl and Jackendoff, 1996; Pankhurst, 2008, pp. 54-55; Schenker, 1954, p. 43). Though Schenker never claimed that surface features are unimportant, his hierarchical removal of elements points to a deeper assumption: without a structured background, the musical surface is somewhat inferior (Schenker's *organic* composition; Hubbs, 1991). The studies presented in this chapter, however, suggest that the musical surface plays a large role in reinforcing the perception of the "structured background"; without these musical elaborations, memory for key would be weakened.

The finding that musical skill did not significantly affect goodness-of-completion ratings is perhaps unexpected. It has been suggested that a participant will pay attention to different elements of music depending on their level of musicianship (Tan & Spackman, 2005). Musicians tend to focus on the structural content of the piece while nonmusicians are more interested in the musical surface (Tan & Spackman, 2005). However, Tan and Spackman (2005) only collected written descriptions of stimuli from the participants. This, coupled with their small sample size (10 musicians, 10 nonmusicians), could have led to a spurious conclusion: it is possible that participants were able to detect similar elements in the music, but skilled musicians had the understanding and the vocabulary necessary to describe them (Cui et al., 2022; Eitan and Granot, 2008; Tillmann and Bigand, 1996; see also Experiment 1 in Deliège et al., 1996). Along this line, McAdams (1989) claimed that the *Ursatz* of the piece is often reflected in the surface features, suggesting that even participants without any musical skill may be able to detect something of the deep harmonic structure of the stimulus (a claim also supported by Cui et al., 2022; Lalitte & Bigand, 2006).

The operationalisation of *Figuration* may be a possible limitation to Experiments 1 and 2: rhythmic figurations specifically (e.g., suspensions and retardations) create off-beat movement. This could give the impression of more rhythmic activity and thus, confound with the factor *Activity*. That said, the instrumentation for stimuli in Experiments 1 and 2 was carefully chosen to address this possible interdependence between factors. A more percussive timbre, such as piano with its swift decay, might have exaggerated this perception; however, sustained bassoon and flute timbres made the duration of notes much clearer. These were used with the expectation that participants would perceive the notes as suspended through time, rather than focusing on onsets only. Experiment 3 did not itself test the difference between stimuli with surface features and homophonic stimuli, relying on previous studies for this comparison. A possible future study could vary these factors and study which has a greater impact on memory decay, but the suspicion would be that homophonic stimuli would lead to faster memory decay trajectories.

3.7 Conclusion

Surface features are commonly considered ornamentations and therefore, less hierarchically significant in music analysis. However, results from the experiments presented in this chapter challenge this assumption by demonstrating that surface features strengthen and expand memory for large-scale tonal structures. As such, these features are vital for our musical experience.

Experiment 3 further demonstrated a clear decay over time, a finding that is contrary to studies supporting cue-based memory models which argue that there is no time-based memory decay. Results instead provide strong support in favour of activation theories and, subsequently, standard models of memory.

Results such as these also challenge the use of homophonic stimuli in musical research as the addition of compositional embellishments (e.g., surface features) may significantly influence results and thus provide a fuller picture of the phenomenon. Furthermore, memory research of large-scale tonal structures is still relatively novel and there is much yet to explore. It has been suggested, for example, that well-formed diatonic excerpts produce significant interference in harmonic memory (Farbood, 2016). However, Experiments 1 through 3 did not exhibit this. How, then, does harmony influence musical memory, if at all?

Chapter 4

Harmony

> Compositions. . . have a point of gravitation, an explicit or implicit center around which all its pitches orbit. This phenomenon is called tonality, and the gravitational center—a single pitch. . . the tonic. (Laitz, 2012, p. 3)

Chapter 3 tested the degree to which surface features influenced perception and memory. The addition of surface features resulted in significantly stronger harmonic memory than previous findings by Farbood (2016) and Woolhouse et al. (2016). However, the intervening sections of all three experiments discussed were composed in a homophonic texture which Farbood claimed did not result in any significant trends. Experiment 3 of Chapter 3 provided evidence that refutes this argument. However, the nature of tonality in the intervening section was not tested. The study reported in this chapter attempts to fill this gap.

Tonality, key and scale are somewhat interdependent; a chromatic composition, for example, may have a scale, but not a strongly perceptible tonal center.

However, a functional diatonic composition—that is, one written using tonal harmonic rules (see Kostka & Payne, 2004; Laitz, 2012)—has both a scale and a clear tonic, both easily perceivable. The scale of a functional diatonic composition follows a pattern of whole and half steps (Kostka & Payne, 2004). This maintains a tonal hierarchy (Krumhansl & Kessler, 1982), the tonic of which, according to Laitz, is its gravitational center. It, therefore, has a high degree of *tonal coherence*. A functional chromatic composition arguably does not follow such a hierarchy, as it may employ equal and balanced step sizes, making it difficult to—perceptually—identify the tonic. However, this arguably has an intermediate degree of tonal coherence, as a *functional* composition, regardless of harmonic hierarchy, must be written following some preconceived compositional pattern and, therefore, must have more cohesion than pure randomness.

For the purposes of this paper, I will use the word *tonal coherence* to represent the degree of harmonic structure from a perceptual standpoint. A functional diatonic composition has high perceptual tonal coherence due to its tonal hierarchy and clear tonic. Functional chromatic compositions are perhaps a step below its diatonic sibling as it has structure, but not such a strong hierarchy nor an easily perceptible tonic. Finally, random compositions of chords occupy the lowest rung of tonal coherence with, perhaps, atonal compositions occupying the space directly above (Lerdahl, 1992). As such, the conditions found in Farbood (2016) can be ranked as such: first, functional diatonic; second, single-chord sequences; and last, random sequences. Her results rank these sequences from single chord, to random, to diatonic last in terms of memory strength. Single-chord trials presumably had least distractibility from the task, the random trials had no clear cadence (which would immediately indicate the key of the sequence), and in last place, functional diatonic sequences—with cadences and a perceptually clear key—were the most interrupting to harmonic memory.

4.1 Experiment 4

This study tested whether the tonal coherence of the intervening section of a non-adjacent key paradigm influenced goodness-of-completion ratings. As such, three main factors were tested: *Nonadjacent Key Relationship* (congruent, incongruent), *Intervening Harmony* (diatonic, chromatic), and *Intervening Chord Pattern* (functional, non-functional).

According to findings from Farbood (2016), and the definitions discussed above, I predicted that goodness-of-completion ratings will be higher for conditions with ambiguous, though organized, non-cadential intervening sections. In other words, the highest ratings will be for the chromatic functional condition, followed by non-functional conditions, and lastly by tonal functional conditions (see Table 4.1).

Table 4.1
Predictions Based on Farbood (2016) Results

GoC	Condition
$\Uparrow$	functional chromatic
$\uparrow$	non-functional chromatic/diatonic
$\downarrow$	functional diatonic

Note. Arrows correspond to the strength and direction of goodness-of-completion (GoC) ratings.

4.1.1 Methods

Participants

Seventy-eight undergraduate students (female: 45, male: 29, nonbinary: 1, undisclosed: 3) participated in this study (ages: 18-34 years; $M = 18.88$, $SD = 2.25$). Twenty participants had five or more years of musical lessons (20 musicians, 58 nonmusicians). A power analysis was conducted for a 2x2x2 repeated measures design using the Superpower package in R (Lakens & Caldwell, 2021), which recommended a minimum of 30 participants. Students received one course credit for participation.

Apparatus

Similar to Experiment 2 in Chapter 3, stimuli were created using *MuseScore3* (MuseScore Project, 2015). The sonority used was the default piano provided by the program as there were no surface features to control for in this study. Tempo was 80bpm. Each stimulus lasted 14 seconds. Stimuli were presented to participants through a program built in *PsychoPy3* and *PsychoJS* (Peirce et al., 2019) and hosted on Pavlovia servers. Due to the COVID-19 pandemic, the study was conducted virtually. As such, ogg files were used to minimize file size and load time for participants. Headphones and devices were out of the control of the experimenters.

Procedure

Informed consent and a basic demographic questionnaire were presented at the beginning of the experiment through *LimeSurvey* (Limesurvey GmbH./LimeSurvey, 2020). Upon completion of the survey, participants were guided to the *PsychoPy* experiment where they were given brief instructions and three novel practice trials. The instructions provided included requests to set aside the appropriate amount of time to run the experiment, to sit in a quiet environment and to use headphones for the experiment. They were further instructed to listen attentively to the musical excerpt and rate the probe (identified by a short pause) on how well it completed the preceding musical phrase. This was clarified as a feeling that the musical excerpt has finished as opposed to feeling that "something was missing". Judgements of completion were rated on a Likert-type 7-point sliding scale (1 = *not at all* and 7 = *strong sense of completion*). Ninety-six (96) stimuli were presented in total, resulting in a 45-minute experiment.

Stimuli

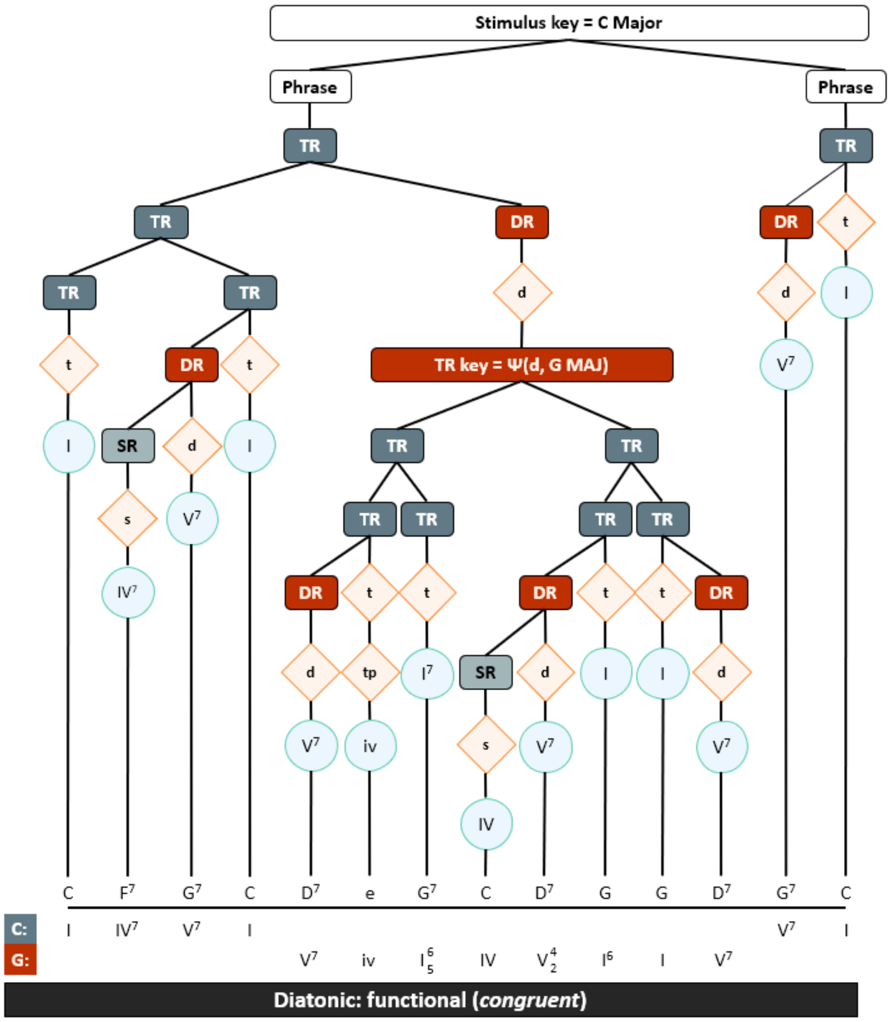
Stimuli were constructed using the nonadjacent paradigm; nonadjacent and probe sections were always written in a well-formed diatonic four-part chord progression in a traditional Western musical style (see Figures A2.1 and 4.4) and could be harmonically congruent (i.e., the stimulus began and ended in the same key X_{ns}-Y_{is}-X_{pc}), or incongruent (Z_{ns}-Y_{is}-X_{pc}). The intervening section was modified so the degree of tonal coherence was varied between trials (see figures below). Consequently, there were three main factors of interest: *Nonadjacent Key Relationship* (congruent or incongruent), *Intervening Harmony* (diatonic or chromatic), and *Intervening Chord Pattern* (functional or non-functional).

A harmonically congruent condition is presented in Figure A2.1. This inverted tree representation, based on Rohrmeier (2011), provides a hierarchical, generative illustration of harmonic progressions that is particularly useful for comparing modulations and differences in compositional decisions between stimuli (compare Figures A2.1, 4.2, and 4.3). The hierarchically highest level indicates the overarching stimulus key, C major in this case. At the second level, the main stimulus and the probe cadence are naturally broken into two distinct phrases by the short pause between Y_{is} and X_{pc}. The intervening section is understood through its relationship to the nonadjacent section (akin to Temperley, 2007 where the order of presentation affects the interpretation of incoming events), as demonstrated on the third and fourth levels. Following levels in the tree are organized in a similar fashion; increasingly specific key regions are constructed from perceptual—and analytical—relationships between chords and/or other key regions. Hierarchically speaking, subdominant regions (SR) prepare dominant regions (DR) which, in turn, prepare tonic regions (TR). Tonic regions are superordinate in the harmonic hierarchy and will always occupy the highest levels of the tree. In Figure A2.1, the stimulus moves from a tonic region (X_{ns} in C major) to a dominant region (Y_{is} in G major) and returns to a tonic region for the probe cadence (X_{pc} in C major).

The incongruent condition modulated to a new key in between every section, as in Experiments 2 and 3 in Chapter 3. Figure 4.2 illustrates this relationship: as the target key of the stimulus, X_{pc} holds the tonic region. The remaining sections (Z_{ns} and Y_{is}) have a dominant relationship to the stimulus key of C major. In this case, Z_{ns} is processed with regards to its relationship to Y_{is} due to the stimulus design (which considers X_{pc} as the 'home' key, Y_{is} as the dominant of that and Z_{ns} as the dominant of the dominant). This is reflected in the subdivision of phrase key

Figure 4.1

Functional Diatonic Condition: Congruent



Note. Key regions are depicted as TR (tonic region, *t* = tonic, *tp* = tonic parallel), DR (dominant region, *d* = dominant), and SR (subdominant region, *s* = subdominant). Roman numeral analysis is divided corresponding to the three parts: nonadjacent (4 chords), intervening (8 chords), and probe (2 chords) sections.

(G major) into a dominant and tonic region. This arrangement has no harmonic return and should, thus, elicit low completion ratings. Modulations were either 1, 3, or 5 steps on the cycle-of-fifths.

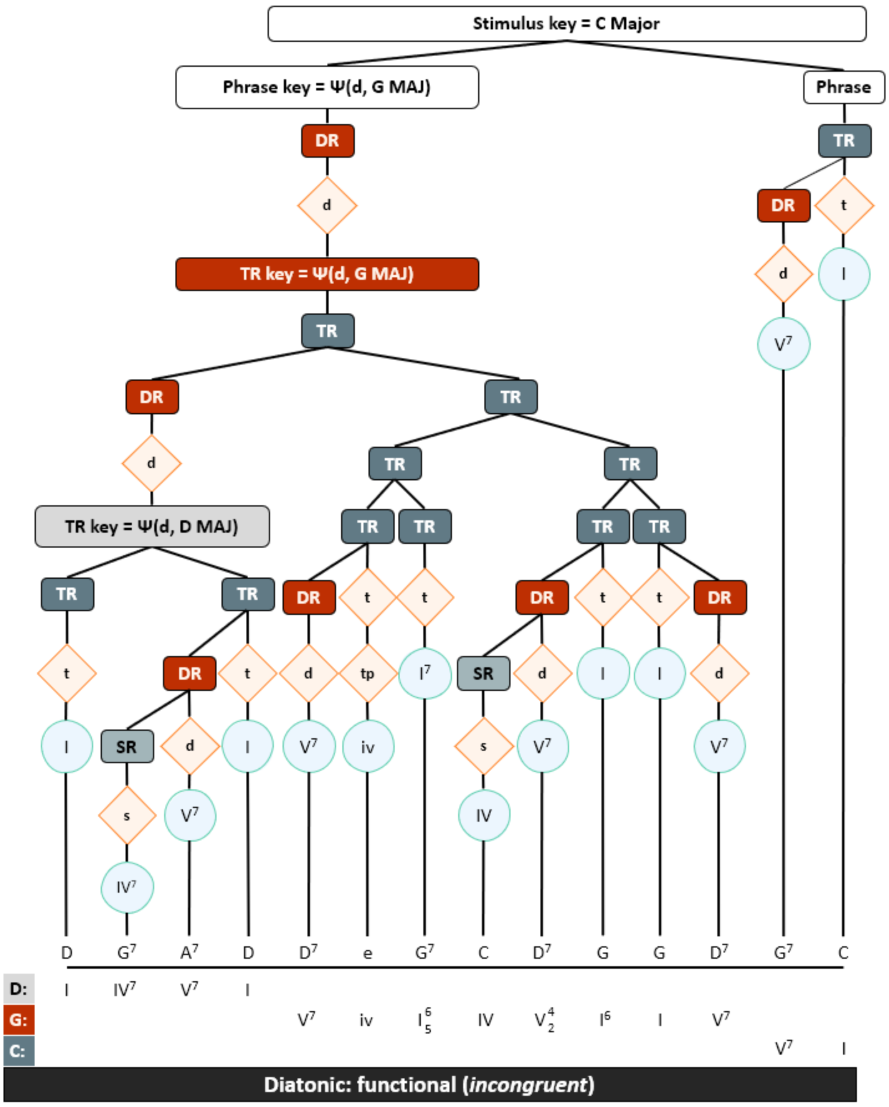
The second factor, *Intervening Harmony*, refers to the intervening section (Y_{is}) of the stimulus which could be either diatonic or chromatic (compare Figure 4.4a and 4.4c). Diatonic stimuli had the most discernable tonal center whereas chromatic stimuli were used to introduce ambiguity to the perceived key of the section.

Factor *Intervening Chord Pattern* describes the chord progression of the intervening section which could be either functional or non-functional. The functional diatonic condition is similar to the other sections of the stimulus, forming a clear representation of the intervening tonal center (see Figures A2.1 and 4.2). Conversely, a non-functional diatonic condition was composed using the same chord progression, but each chord was quasi-randomly transposed up or down 1-6 semitones such that no repetitions occurred. This created a sequence which had no perceivable tonal center (see Figures 4.3 and 4.4). Figure 4.3 demonstrates the randomization process: the stimulus begins and ends in C major, a congruent nonadjacent key relationship (though incongruent relationships were also used), chords remained the same as in their functional counterparts, but randomly shifted up or down. As such, a IV chord in C major, transposed by a random number, will become IV of a different key. The corresponding keys are indicated above the analysis (in orange).

A similar method was used for the chromatic versions of these factor levels; a functional version was composed to match as closely as possible to the diatonic

Figure 4.2

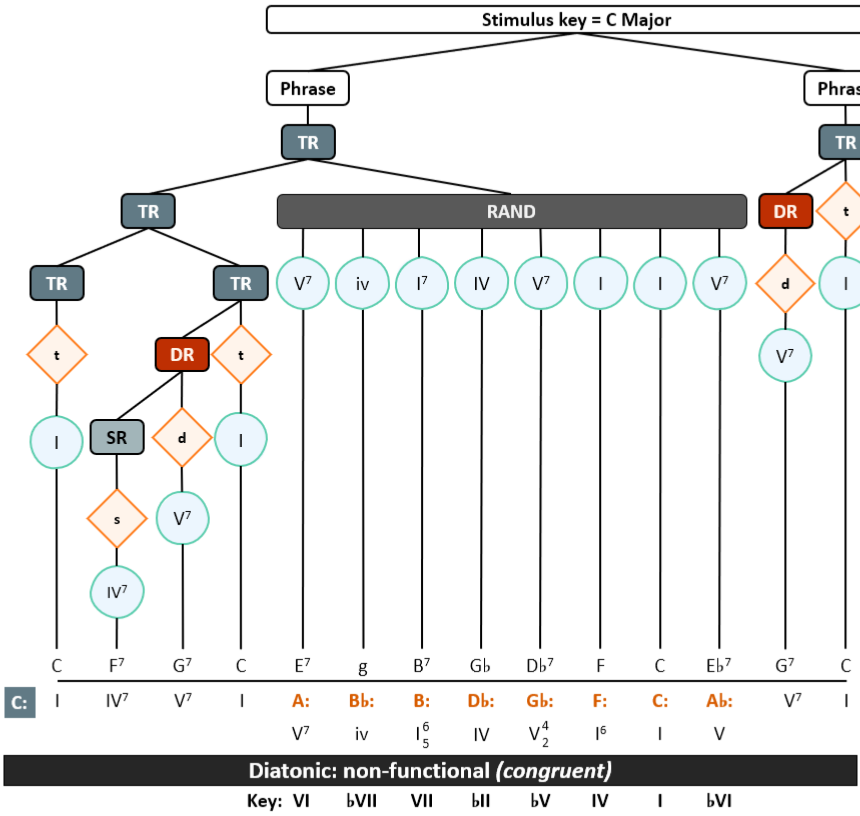
Functional Diatonic Condition: Incongruent



Note. Roman numeral analysis is now divided into three levels, pertaining to each change in key.

Figure 4.3

Non-functional Diatonic Condition



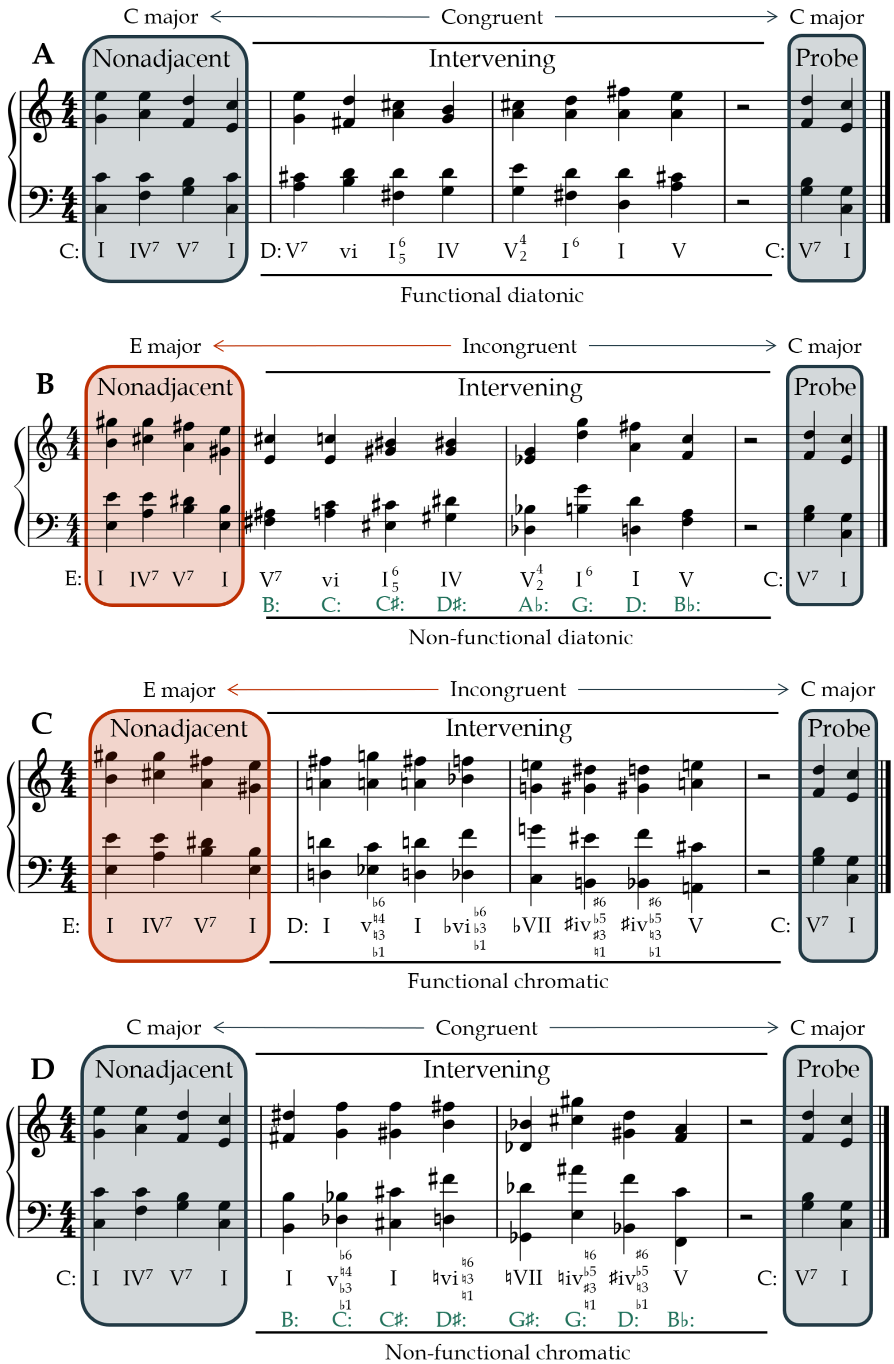
Note. Keys are once again indicated below the figure (in black) but now in relation to the stimulus key (C major) using roman numeral analysis.

functional stimulus in terms of number of dissonant and consonant chords and their placement within the sequence. The non-functional condition was quasi-randomly modulated as described above (see Figure 4.4).

Control variables included the *Modulation Distance* (1, 3, and 5 steps on the cycle-of-fifths), *Modulation Direction* (up, down), and *Probe Sequence Key* (any major key). Each condition (e.g., congruent chromatic non-functional) was presented in all modulation distances and directions forming 48 combinations. Keys were assigned quasi-randomly to each stimulus probe section and modulations were

Figure 4.4

Samples of all Conditions



Note. Sample stimuli comparing (A) a congruent functional diatonic condition, (B) an incongruent non-functional diatonic condition with the key corresponding to each chord indicated below the roman numeral analysis, (C) an incongruent functional chromatic condition, and (D) a congruent non-functional chromatic condition.

organized backwards such that stimuli ended on these predetermined keys. The stimulus set was presented twice to each participant, with randomized probe keys (resulting in 96 stimuli in total).

4.1.2 Results

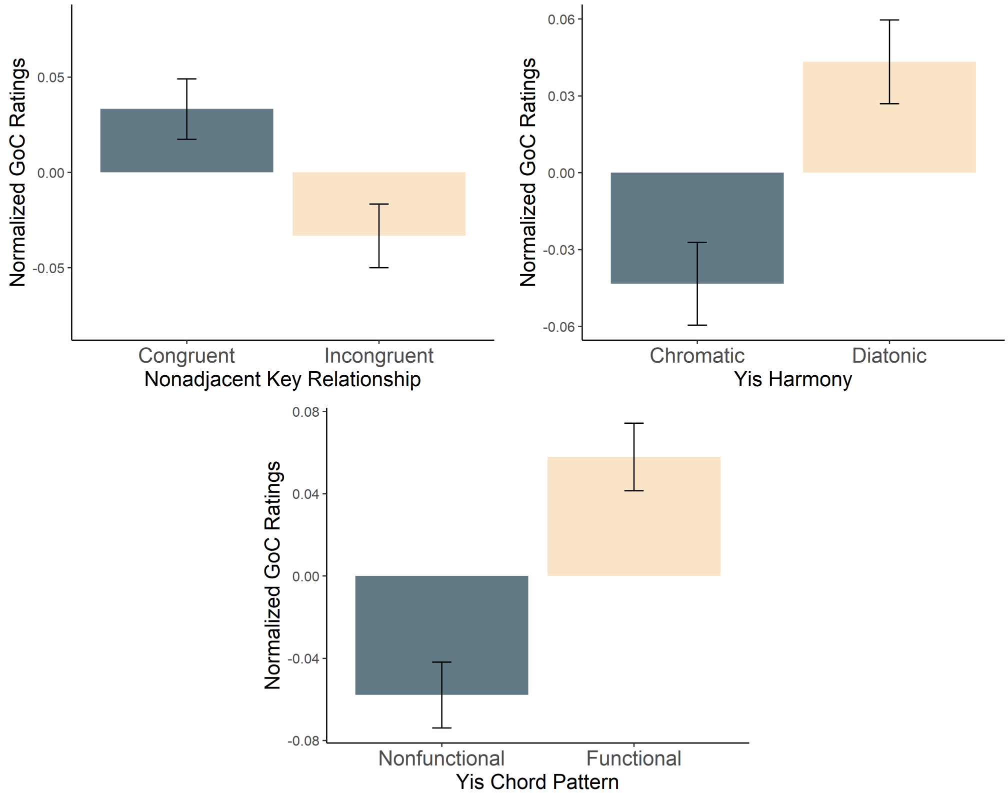
Ratings were averaged across all control variables such that only the conditions pertaining to the three main factors remained. A 2x2x2 repeated measures ANOVA was conducted with factors *Nonadjacent Key Relationship* (congruent, incongruent), *Intervening Harmony* (diatonic, chromatic) and *Intervening Chord Pattern* (functional, non-functional). Main effects were all significant. Factor *Nonadjacent Key Relationship* ($F_{1,\,77} = 7.69$, $p < 0.01$, $\eta_p^2 = 0.09$) was rated higher for congruent key relationships than incongruent ones ($M_{\text{incongruent}}$ = -0.033, $M_{\text{congruent}}$ = 0.033). *Intervening Harmony* ($F_{1,\,77} = 9.98$, $p < 0.01$, $\eta_p^2 = 0.12$) was rated higher for conditions where Y_{is} was diatonic rather than chromatic ($M_{\text{chromatic}}$ = -0.043, M_{diatonic} = 0.043). Lastly, *Intervening Chord Pattern* ($F_{1,\,77} = 10.57$, $p < 0.01$, η_p^2 = 0.12) was rated higher for functional than non-functional stimuli ($M_{\text{nonfunctional}}$ = -0.058, $M_{\text{functional}}$ = 0.058; Figure 4.5).[1]

There was a significant interaction between *Intervening Harmony* and *Intervening Chord Pattern* ($F_{1,\,77} = 5.94$, $p < 0.05$, $\eta_p^2 = 0.07$) where functional diatonic stimuli were rated significantly higher than any other condition (Functional: M_{diatonic} = 0.135, $M_{\text{chromatic}}$ = -0.019; Nonfunctional: M_{diatonic} = -0.048, $M_{\text{chromatic}}$

[1] As in the experiments in Chapter 3, normalized scores were used for visualization purposes only. Analyses were all done using raw data.

Figure 4.5

Main Effects: Nonadjacent Key Relationship, Harmony, and Chord Pattern

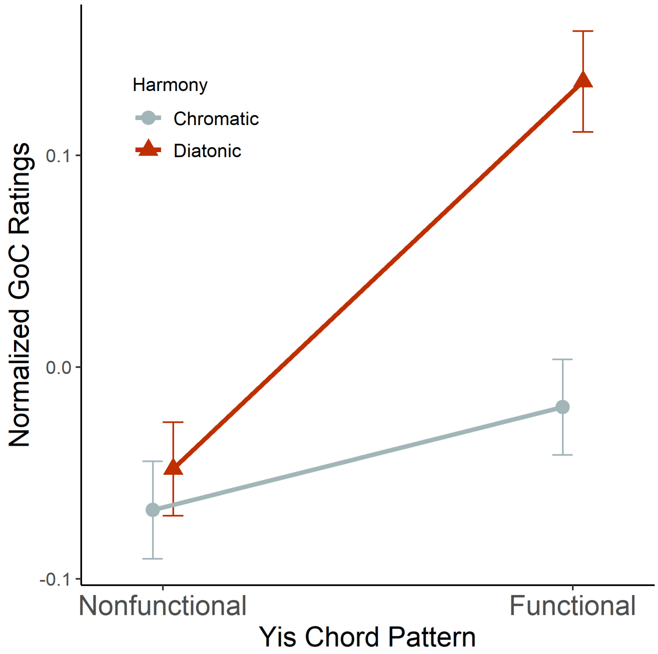


Note. Mean ratings across participants, with 95% confidence intervals.

= -0.068). Functional chromatic stimuli seemed to be rated higher than non-functional conditions but did not reach significance in a post-hoc Tukeys HSD (see Figure 4.6).

The interactions between *Nonadjacent Key Relationship* and *Intervening Harmony* ($p = 0.97$) and *Nonadjacent Key Relationship* and *Intervening Chord Pattern* ($p = 0.74$) did not reach significance. The 3-way interaction between *Nonadjacent Key Relationship*, *Intervening Harmony*, and *Intervening Chord Pattern* also did not reach significance ($p = 0.48$).

Figure 4.6
Interaction: Chord Pattern and Harmony



As in Chapter 3, factor *Musician* was added in post-hoc to the repeated-measures ANOVA but did not interact with any independent variable or interaction described above.

4.1.3 Discussion

Farbood (2016) found large differences between stimuli with functional diatonic, non-functional diatonic and single chord intervening sections where, notably, the functional diatonic condition led to rapid loss in memory of the nonadjacent key whereas the single-chord condition resulted in prolonged harmonic memory. A possible limitation, therefore, with the three experiments discussed in Chapter 3 was that the intervening section was always composed in a functional diatonic style.

However, the experiments in Chapter 3 found encouraging results: tonal memory extended past even Farbood's 21 second limit. If participants were indeed influenced by the intervening tonality, arguably there should have been little to no significance in any of the three studies. The experiment discussed in this chapter addressed this discrepancy by testing intervening harmony in four levels: functional diatonic, non-functional diatonic, functional chromatic, and non-functional chromatic, each with differing degrees of *tonal coherence*. Along with significant main effects, there was a significant interaction between *Intervening Harmony* and *Chord Pattern* in which functional diatonic stimuli were rated significantly higher in completion than any other condition. This directly contradicts the findings in Farbood (2016).

One possible explanation for this discrepancy is that the task itself was different. Subjective ratings of tension or goodness-of-completion both have different temporal perspectives. Tension ratings focus on making judgements in the present moment, but completion ratings are retrospective in nature; participants must reflect on the entire stimulus in order to form the judgement. It is possible that this difference could be influencing ratings differently.

Second, the nature of both studies was strikingly different: Farbood ran the experiments in person in a laboratory whereas the experiment described in this chapter was completely virtual. Each would have brought different extraneous variables into the fold. The unnatural nature of doing an experiment in a scientific laboratory brings a specific kind of anxiety. Conversely, despite the comfort of the home, completing an experiment amid a global pandemic is perhaps also stressful in its own way. Two of the three experiments described in Chapter 3, however, were run in person before the pandemic began. The results from both strongly support the findings of this chapter, most notably Experiment 3 which extended memory decay past the limits discussed in Farbood (2016). To conclude, though it is not unlikely that the virtual nature of the experiment introduced noise to the results, converging evidence from four experiments run both virtually and in person strongly suggests that the use of a functional diatonic sequence in the intervening section is valid and does not disrupt harmonic memory.

A possible limitation of the current study is that single-chord trials were not tested. These were significantly more successful in Farbood (2016) and, as such, it is possible that a single-chord condition could have boosted memory even further than what was seen in the experiments in Chapters 3 and 4. However, current results show no evidence to support this hypothesis. Furthermore, Farbood described the participants' boredom in having to spend over 20 seconds listening to the same chord and expressed their relief at finally hearing a cadence after this monotony. This certainly could have influenced participant data. Perhaps with shorter stimuli, this could be a useful factor to test in conjunction with the nonadjacent relationship paradigm and surface features.

4.1.4 Conclusion

Previous research has suggested that the functional diatonic composition of the intervening section of a nonadjacent key paradigm provides a high degree of interference in harmonic memory (Farbood, 2016). However, results from the experiment described in this chapter suggest that this is not the case. In fact, the opposite was found to be true: the functional diatonic condition was the only one to bare significant difference in goodness-of-completion ratings. These results are further supported by the findings of experiments discussed in Chapter 3, as well as from Spyra et al. (2021) and Woolhouse et al. (2016). The converging evidence from these studies suggest that memory is not impaired by functional tonality and may, in fact, be bolstered by it.

Chapter 5

Following Perception to Judgement

> ...in the process of amalgamating, the dependent parts give up, to some degree, the properties they had as individuals. (Bregman, 1994, pp. 474-5)

As set out at the start of this thesis, memory for large-scale tonal structures is not often studied. This line of research began with Cook in 1987 who tested preference for modulating and non-modulating versions of piano excerpts. He found significant differences only for excerpts lasting up to one minute in length. However, his measures lacked specificity; he did not compare lengths of modulations in his stimuli, leaving the reader unsure whether the differences in judgement were due to memory for the original key, the local effect of switching keys, or some other hidden factor. This study was followed by Woolhouse et al. (2016) who tested memory for large-scale tonal structures through a nonadjacent key paradigm. This paradigm allowed them to fill the gap Cook (1987) left. Their study found that memory for a key had decayed rapidly by approximately 11 seconds regardless of musical experience. Farbood (2016) expanded on this research by modifying the intervening section of the nonadjacent key paradigm by degree of tonal ambiguity in addition to duration. She found no significant effects in functional diatonic conditions but significantly extended effects in a single-chord condition. This condition pushed the limit of memory decay to 21 seconds. Spyra et al. (2021) continued this work by testing whether memory did, indeed, decay over time as Farbood (2016) and Woolhouse et al. (2016) had claimed or whether the number of intervening events replaced the original key in memory instead. They found significant results for time decay but no significant differences for number of chords/events providing support to the findings of the previous researchers.

Key discoveries from these studies tell us that: (1) tonal memory lasts in the seconds, as opposed to hours or days as would be expected from instances of remarkable memory in previous research, (2) memory decays rapidly over time, (3) it is not interrupted by intervening information, and (4) it is significantly influenced by tonal features of intervening information. Though Farbood (2016), Spyra et al. (2021), and Woolhouse et al. (2016) all provided excellent foundational work on memory for large-scale tonal structures, the issue remains that these stimuli may not generalize to real-world music; people rarely listen to only homophonic music of this degree of simplicity.

The experiments in Chapter 3 addressed this concern by introducing surface features, in the form of melodic and rhythmic embellishments, to nonadjacent sections. These experiments found that: (1) surface features increased the sense of structural cohesion of the stimuli, (2) there was an increase in harmonic memory when surface features were present, and (3) that surface features extended the duration of harmonic memory to approximately 33 seconds after modulation, well past that predicted by Farbood (2016) and Woolhouse et al. (2016).

However, these results are somewhat contrary to those in Farbood (2016) as the intervening sections in these experiments were highly tonal. Farbood found that functional diatonic compositions of the intervening section resulted in insignificant tension ratings, suggesting that memory for the original key decayed more rapidly after modulation to a highly tonal intervening key than after other less tonal-functional conditions. In response, the study described in Chapter 4 systematically varied the degree of tonal coherence in the intervening section. It was found that, of all conditions, the functional diatonic alone differentiated significantly from the rest. These contrary results were further supported by the significance and effect sizes of experiments in Chapter 3.

A variety of factors could, possibly, account for this discrepancy. The task itself was different: Farbood measured tension whereas my study used completion ratings for the probe cadence. It is, therefore, not unlikely that both studies measured memory for the original key in different ways. Most notably is that the perspective is different: tension ratings measure memory continuously—in the present—whereas completion ratings are retrospective. This difference in view could be targeting different aspects of the memory process. The nonadjacent-key paradigm needed to activate working memory, for example, whereas a continuous tension rating measures memory in an unconscious form. Further research would be needed to clarify this in more detail.

The presentation of the experiments was also different. Farbood ran her experiments in person before the pandemic arrived, whereas two of the four experiments described in Chapters 3 and 4 were virtual. That being said, Experiments 1 and 3 (Chapter 3) were run in the laboratory before the COVID-19 pandemic with compelling results. Experiment 3, in particular, demonstrated the significant increase in the duration of memory decay from Farbood's 21-second limit. Though we may certainly question the generalizability of data collected during a global pandemic, the results gathered in person provide substantial and compelling support for those gathered virtually.

A possible limitation of the study presented in Chapter 4 is the fact that single-chord conditions were not tested. This was mainly due to time constraints and the systematic nature of the manipulations. However, the significant and clear difference between the functional diatonic condition, the random conditions (which Farbood also employed) and the functional chromatic condition provides overwhelming support for the hypothesis that diatonic intervening sections do not, in fact, interfere with harmonic memory. Though this is contradictory to the findings in Farbood (2016), it is supported by the three experiments in Chapter 3—which should not have had significant results according to the trends found in Farbood (2016)—and by Spyra et al. (2021) and Woolhouse et al. (2016) who found significant results despite functional diatonic intervening sections. Without explicitly testing the single-chord condition, however, we cannot accurately predict whether a single-chord condition might have further boosted harmonic memory even in a goodness-of-completion measure. Further testing must be done to illuminate the issue.

5.1 Music and Memory

The study of memory began with Ebbinghaus in the 1880's. Since then, it has evolved in complexity through James' primary and secondary memory (James, 1950), was further subdivided in the modal model by Atkinson and Shiffrin (1968), and subdivided again by Baddeley and Hitch (1974) in their model of working memory. In short, through 100 years and more, there was an ever-increasing specificity in the way science thought about the processes involved in memory.

Chapter 1 explored the memory process starting from echoic memory and the psychoacoustic model by Leman (2000) who mapped the progress of memory decay from the earliest stages of the process. Echoic, short-term, and long-term memories all have unique capacities and limits through varying decay rates over time. Baddeley and Hitch (1974) explored working memory (WM), which is often described as a part of the short-term memory (STM) process. They divided WM into various rehearsal loops according to specific domains (e.g., visual, auditory, computational, etc.). However, Baddeley and Hitch's phonological loop was challenged by music researchers, who argued that music and language have their own rehearsal loops, if perhaps overlapping somewhat in neural resources when processing syntax (Fiveash & Pammer, 2014). Conversely, Cowan (2008) argued that all memory is one: WM is simply a subset of activated STM, which in turn is a subset of activated long-term memory (LTM). A common theme between these models, including Cowans model, is the argument that memory decays over time.

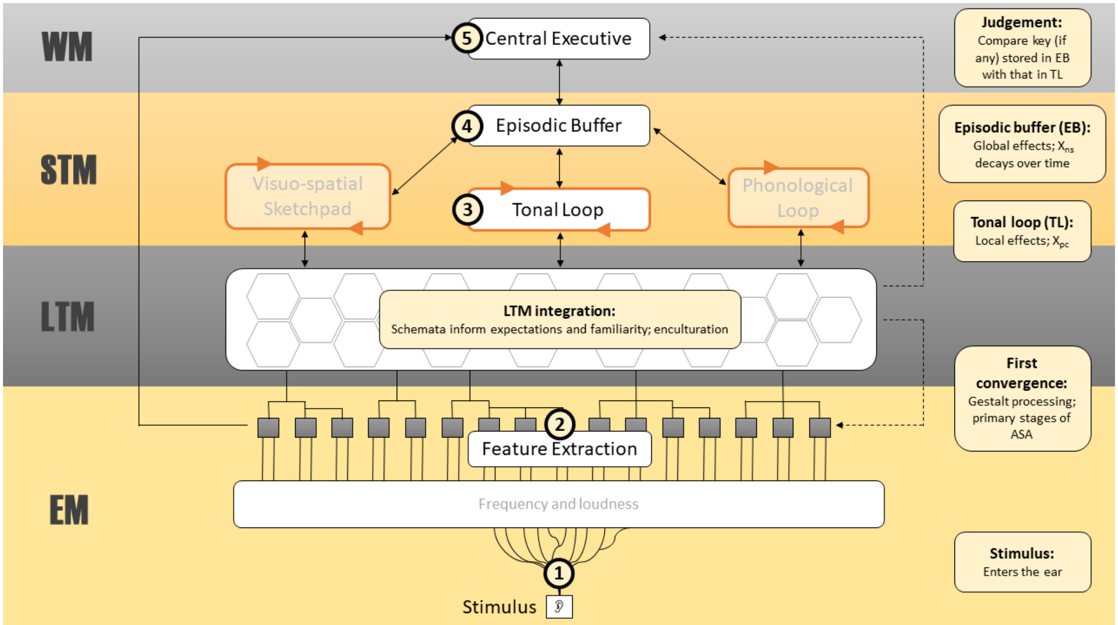
Not all researchers agree that memory decays over time, however. Alternative models, such as the Feature Model (Nairne, 1990) and OSCAR (Brown et al., 2000), rely on processes other than decay to explain forgetting. Nairne's Feature Model argues that an increase of cues dilutes the extent to which the target memory can correctly be identified. The OSCAR model itself does not have a component that would allow for a decay mechanism. Once activated, its oscillators would keep running at the same frequency indefinitely or until replaced by another item. Again, cues might become confused if the number of items to recall is great, but time alone cannot degrade the reactivation of the correct memory. Furthermore, Lewandowsky et al. (2004) have directly tested memory decay and found no evidence to support such a mechanism.

The work by Farbood (2016), Spyra et al. (2021), and Woolhouse et al. (2016), as well as Experiment 3 in Chapter 3, as described above, are not consistent with this view. All researchers found clear examples of memory decay in their experiments using the nonadjacent key paradigm. The Feature Model in particular would have predicted that the number of events in Spyra et al. (2021) would lead to significant results rather than duration; specifically, six intervening events should have interfered with harmonic memory more than four events did as it provides more cues. Conversely, if six events/chords were not enough to weaken the link to the original key, perhaps a null result overall. This was not found, however.

Given this, and issues arising from standard models (Nairne, 1990), I have proposed a cognitive system that incorporates models and understanding by Baddeley and Hitch (2019), Cowan (2008), Leman (2000), and Snyder (2000), as well as evidence from music cognition experiments such as that by Berz (1995). The benefit of this system is that one could follow a judgement task—such as those used in Experiments 1-4—from sound to resolution. The idea of the task fitting neatly into just one of several memory components is naïve; a judgement task such as this uses *every* part of the system as demonstrated below.

Figure 5.1

Model of Memory



Note. My proposed model, including in it echoic memory (EM), long-term memory (LTM), short-term memory (STM) and working memory (WM), from Chapter 1, p. 10.

Chords first enter through the sensory organ (Figure 5.1, bottom), vibrating tympanic membranes, activating hair cells, and making their way up through to the brain stem. EM is the first process they encounter; here, basic features such as frequency and loudness of the chords are parsed (Snyder, 2000, p. 4). Importantly, there is no direct connection to LTM at this stage and, therefore, no labeling can occur at this point (i.e., there is no knowledge of what kind of pitch is heard, only the general features of it is known; Figure 5.1, first layer near bottom). EM has a brief but large storage capacity and much information about the sound of our chords can be kept here. This raw information converges in the *feature* *extraction* stage (Snyder, 2000, p. 7). At this stage, features bind into timbres, notes, and other musical elements, though still with none to limited labeling from LTM (Figure 5.1, top layer of grey squares in EM). Here, we hear the difference between a bassoon, flute, or piano timbre from Chapter 3. Processing thus far has been based on Gestalt-like processing found in auditory scene analysis (Alain & Bernstein, 2008; Bendixen, 2014; Denham & Winkler, 2006; Rogers & Bregman, 1993).

Feature bundles connect directly to LTM and STM. LTM feeds back into feature extraction in a subtle but important way: based on information stored in LTM, we can form a *context* for the current experience (Snyder, 2000, p. 5). LTM is where veridical and schematic knowledge reside; veridical knowledge—a memory of that exact stimulus—is preferably not being activated in a nonadjacent key paradigm, but schematic knowledge is highly utilized to understand common musical conventions, timbres, and the nonadjacent keys that are compared when we make a completion judgement. Our schemata help us understand a cadence and extract important *tonal coherence* (Chapter 4) information from it.

Meanwhile, the sequence of chords enters STM where the excerpt is temporarily stored in the tonal loop. This, too, is informed by schemata in LTM. The sequences rehearsed here are responsible for *local* effects: the way we perceive music as it modulates and builds on itself. However, as discussed in Chapter 1, certain stimuli require additional STM resources. By the time the excerpt moves into X_{pc}, the nonadjacent harmony has relocated to the episodic buffer. This buffer is what is responsible for *global* effects; the interpretation of current musical events is often coloured by musical information gathered in the—relatively—recent past. Both buffer and tonal loop are required for the final goodness-of-completion judgement to successfully take place. The central executive—the center of attention and working memory—compares the key that is stored in the episodic buffer with the key currently residing in the tonal loop and makes a judgement: are they the same or, if not, how dissimilar are they? Unless the duration of the intervening section has passed some threshold, the key residing in the buffer is X_{ns} or Z_{ns}. If, however, enough time has passed, that key has been replaced by Y_{is} and the task can no longer be completed successfully. Whether informed by the nonadjacent harmony or not, a judgement is made, and the process begins anew.

5.2 Conclusion

The research described in these chapters is uniquely important as it illustrates a process that is little understood in music cognition research. Memory involving harmonic structures is rarely addressed, yet large-scale tonal structures are used so frequently in music. Composers often rely on keys for instilling a sense of return, or a holistic cohesiveness, to the work. Unfortunately, it seems that human memory for large-scale harmonic structures is relatively poor, spanning less than one minute before it decays (Farbood, 2016; Spyra et al., 2021; Woolhouse et al., 2016). The elements that make up the musical experience, however, are only beginning to be explored and, as experiments on surface features have reflected, these may yet have a significant boosting effect.

Notably, surface features have been historically regarded as mere elaborations in music analytical spheres. Often favoring the *Ursatz*, surface features such as those employed by experiments in Chapter 3 are considered hierarchically less important. However, results here show that this is not necessarily the case in perception and memory and the surface is, in fact, vital to the musical experience of harmonic structure. Though researchers have claimed that the Ursatz may be reflected in surface features, thereby supporting the idea that the surface holds valuable information (Cui et al., 2022; Lalitte & Bigand, 2006; McAdams, 1989), this view may yet be too mild to explain these findings. Results such as these must, therefore, inform music theorists and composers as their most common-place assumptions may, in fact, lack validity.

Findings from Chapters 3 and 4 show that the compositional elements music cognition researchers put into their stimuli matter a great deal. The presence of certain musical elements can give drastically different results and, if not used thoughtfully, could provide spurious results or no results at all. As such, it may be that the effect is not missing, but the stimuli used—such as homophonic stimuli—are too simple to tell the whole story. Conversely, stimuli could be too complex and overcomplicate the story. As discussed in Chapter 4, many elements in music have interdependencies within them; the definition of a key or a melody may seem simple on the surface, but in fact correlate with scales, pitch height and many other compositional dimensions. A music cognition researcher must, therefore, be cognizant of these interdependencies and select/create stimuli carefully.

In summary, this research is unique. Many studies in music cognition focus on longer lasting effects, such as reminiscence bumps (Krumhansl, 2017; Krumhansl & Zupnick, 2013), that are often astonishing examples of human memory. However, large-scale tonal structures are contradictory to these examples. They reside in a type of memory that is active for less than a minute (STM/WM). Compared to a lifetime of learning and building of schemas, this is fleeting indeed. Perhaps the model in Figure 5.1 can help elucidate this difference a little. Schemata and veridical memories from the reminiscence bump are both long-term processes. However, much more research is needed to understand the phenomenon of the nonadjacent key relationship which uses schemata in processing presented information, but resides in many other areas of the model as well. The future of this research is vast and bright as it can provide some answers to what makes memory for music so robust and versatile, and yet, fleeting.

Appendix A

Chapter 3 Supplement

Supplemental material for linear mixed effects models in Experiment 3, outlining the models used, the corresponding ANOVA, and a further breakdown of results per model.

Table A1.1
Linear Mixed Effects Models

Models:
M1: Response 1 + (1|Participant)
M2: Response 1 + *Yis Duration* + (1|Participant)
M3: Response 1 + *Yis Duration* + *Nonadjacent Key Relationship* + (1|Participant)
M4: Response 1 + *Yis Duration* + *Nonadjacent Key Relationship* + *Yis Duration:NKR* + (1|Participant)

Table A1.2
Corresponding Results: ANOVA

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
M1	3	12721	12740	-6357.6	12715			
M2	4	12598	12623	-6295.2	12590	124.68	1	< 2e-16 ***
M3	5	12596	12627	-6293.2	12586	4.06	1	0.0439 *
M4	6	12595	12632	-6291.7	12583	2.95	1	0.0861 .

*** $p < 0.0001$; ** $p < 0.01$; * $p < 0.05$; . $p < 0.1$

Table A1.3
Further Breakdown of Components: Y_{is} Duration and Nonadjacent Key Relationships (NKR)

	Model 1	**Model 2**	**Model 3**	**Model 4**
Intercept:	3.71 ***	4.32 ***	4.27 ***	4.18 ***
(Duration 0s, NKR incongruent)	(0.07)	(0.09)	(0.09)	(0.11)
Y_{is} Duration mean		-0.03 ***	-0.03 ***	-0.02 ***
		(0.00)	(0.00)	(0.00)
NKR congruent			0.11 *	0.29 *
			(0.05)	(0.12)
Duration * NKR congruent				-0.01
				(0.01)
AIC	12721.12	12598.44	12596.38	12595.43
BIC	12739.46	12622.90	12626.69	12632.13
Log Likelihood	-6357.56	-6295.22	-6293.19	-6291.72
Num. observations	3348	3348	3348	3348
Num groups: Participant	93	93	93	93
Var: Participant (intercept)	0.40	0.40	0.40	0.40
Var: Residual	2.48	2.38	2.38	2.38

*** $p < 0.0001$; ** $p < 0.01$; * $p < 0.05$

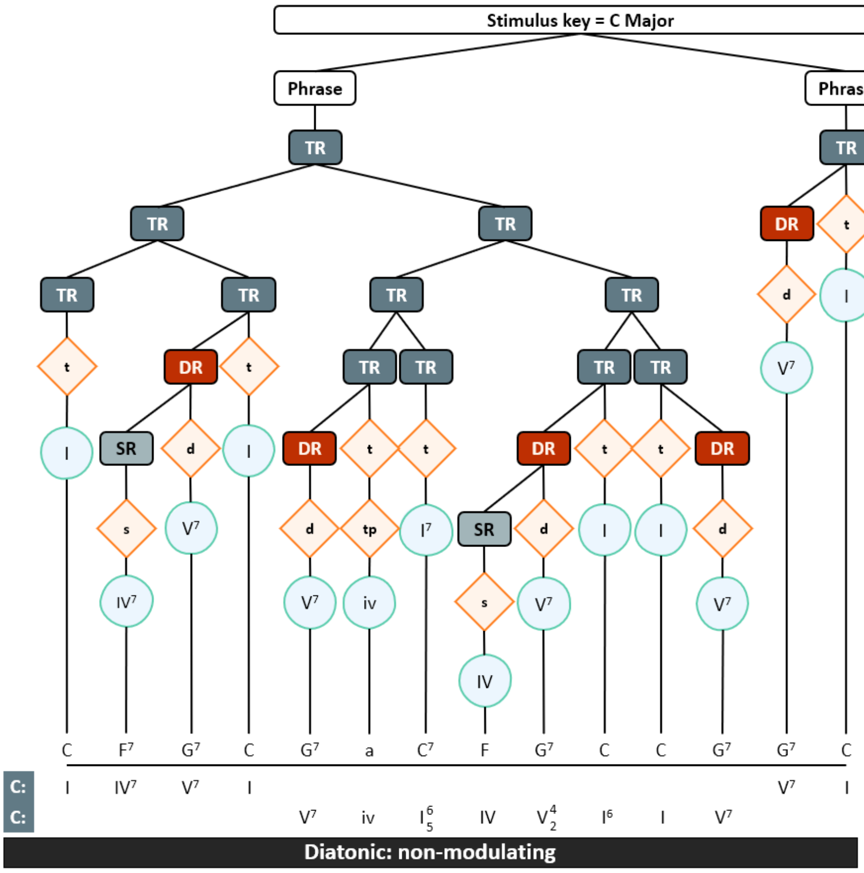
Appendix B

Chapter 4 Supplement

Supplemental material for the harmony trees included in Chapter 4.

Figure A2.1

Functional Diatonic Condition: Nonmodulating



Note. This is the baseline stimulus from which all subsequent stimuli were created.

Bibliography

Agres, K. R. (2019). Change detection and schematic processing in music. *Psychology of Music*, *47*(2), 173–193. https://doi.org/https://doi.org/10.1177/0305735617751249

Akiva-Kabiri, L., Vecchi, T., Granot, R., Basso, D., & Schön, D. (2009). Memory for tonal pitches: A music-length effect hypothesis. *Annals of the New York Academy of Sciences*, *1169*(1), 266–269. https://doi.org/10.1111/j.1749-6632.2009.04787.x

Alain, C., & Bernstein, L. J. (2008). From sounds to meaning: The role of attention during auditory scene analysis: *Current Opinion in Otolaryngology & Head and Neck Surgery*, *16*(5), 485–489. https://doi.org/10.1097/MOO.0b013e32830e2096

Aldwell, E., Schachter, C., & Cadwallader, A. (2010). *Harmony and voice leading* (4th, illustrated ed.). Cengage Learning.

Atkinson, R., & Shiffrin, R. (1968). Human memory: A proposed system and its control processes. *Psychology of learning and motivation* (89–195). Elsevier. https://doi.org/10.1016/S0079-7421(08)60422-3

Auhagen, W., & Vos, P. G. (2000). Experimental methods in tonality induction research: A review. *Music Perception*, *17*(4), 417–436. https://doi.org/10.2307/40285827

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

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*(3), 189–208. https://doi.org/10.1016/S0021-9924(03)00019-4

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*(1), 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Baddeley, A., Allen, R. J., & Hitch, G. J. (2010). Investigating the episodic buffer. *Psychologica Belgica*, *50*(3), 223. https://doi.org/10.5334/pb-50-3-4-223

Baddeley, A., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.). Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1

Baddeley, A. D., & Hitch, G. J. (2019). The phonological loop as a buffer store: An update. *Cortex*, *112*, 91–106. https://doi.org/10.1016/j.cortex.2018.05.015

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4 [preprint arXiv:1406.5823]. *arXiv*.

Bendixen, A. (2014). Predictability effects in auditory scene analysis: A review. *Frontiers in Neuroscience*, *8*. https://doi.org/10.3389/fnins.2014.00060

Berz, W. L. (1995). Working memory in music: A theoretical model. *Music Perception: An Interdisciplinary Journal*, *12*(3), 353–364. https://doi.org/10.2307/40286188

Bey, C., & McAdams, S. (2002). Schema-based processing in auditory scene analysis. *Perception & Psychophysics*, *64*(5), 844–854. https://doi.org/10.3758/BF03194750

Bharucha, J. J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception*, *5*(1), 1–30. https://doi.org/10.2307/40285384

Bharucha, J. J. (1994). Tonality and expectation. *Musical perceptions* (213–239). Oxford University Press. https://doi.org/10.1037/14171-019

Bigand, E., & Parncutt, R. (1999). Perceiving musical tension in long chord sequences. *Psychological Research*, *62*(4), 237–254. https://doi.org/10.1007/s004260050053

Bregman, A. S. (1994). *Auditory scene analysis: The perceptual organization of sound* (1st ed.). MIT Press.

Brown, G. D. A., Hulme, C., & Preece, T. (2000). Oscillator-based memory for serial order. *Psychological Review*, *107*(1), 127–181. https://doi.org/10.1D37//0033-295X.107.1.127

Chen, Z., & Cowan, N. (2005). Chunk limits and length limits in immediate recall: A reconciliation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(6), 1235–1249. https://doi.org/10.1037/0278-7393.31.6.1235

Conway, A. R., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, *7*(12), 547–552. https://doi.org/10.1016/j.tics.2003.10.005

Cook, N. (1987). The perception of large-scale tonal closure. *Music Perception: An Interdisciplinary Journal*, *5*(2), 197–205. https://doi.org/10.2307/40285392

Cowan, N., & Aubuchon, A. M. (2008). Short-term memory loss over time without retroactive stimulus interference. *Psychonomic Bulletin & Review*, *15*(1), 230–235. https://doi.org/10.3758/PBR.15.1.230

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? In W. S. Sossin, J.-C. Lacaille, V. F. Castellucci, & S. Belleville (Eds.), *Essence of memory* (323–338). Elsevier. https://doi.org/10.1016/S0079-6123(07)00020-9

Cowan, N., Saults, J. S., & Blume, C. L. (2014). Central and peripheral components of working memory storage. *Journal of Experimental Psychology: General*, *143*(5), 1806–1836. https://doi.org/10.1037/a0036814

Cowles, A., Beatty, W. W., Nixon, S. J., Lutz, L. J., Paulk, J., Paulk, K., & Ross, E. D. (2003). Musical skill in dementia: A violinist presumed to have Alzheimer's Disease learns to play a new song. *Neurocase*, *9*(6), 493–503. https://doi.org/10.1076/neur.9.6.493.29378

Craik, F. I. (2020). Remembering: An activity of mind and brain. *Annual Review of Psychology*, *71*(1), 1–24. https://doi.org/10.1146/annurev-psych-010419-051027

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Cui, A., Troje, N. F., & Cuddy, L. L. (2022). Electrophysiological and behavioral indicators of musical knowledge about unfamiliar music. *Scientific Reports*, *12*(1), 441. https://doi.org/10.1038/s41598-021-04211-w

Deliège, I., Mélen, M., Stammers, D., & Cross, I. (1996). Musical schemata in real-time listening to a piece of music. *Music Perception*, *14*(2), 117–159. https://doi.org/10.2307/40285715

Denham, S., & Winkler, I. (2006). The role of predictive models in the formation of auditory streams. *Journal of Physiology-Paris*, *100*(1), 154–170. https://doi.org/10.1016/j.jphysparis.2006.09.012

Deutsch, D. (1970). Tones and numbers: Specificity of interference in immediate memory. *Science*, *168*(3939), 1604–1605. https://doi.org/10.1126/science.168.3939.1604

Duncan, M., & Lewandowsky, S. (2005). The time course of response suppression: No evidence for a gradual release from inhibition. *Memory*, *13*(3), 236–246. https://doi.org/10.1080/09658210344000233

Eitan, Z., & Granot, R. Y. (2008). Growing oranges on Mozart's apple tree: "Inner form" and aesthetic judgment. *Music Perception*, *25*(5), 397–418. https://doi.org/10.1525/mp.2008.25.5.397

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. Academic Press. https://doi.org/10.1016/S0079-7421(03)44005-X

Farbood, M. M. (2016). Memory of a tonal center after modulation. *Music Perception: An Interdisciplinary Journal*, *34*(1), 71–93. https://doi.org/10.1525/mp.2016.34.1.71

Fiquet, P., Wu, A., & Hyafil, A. (2021). *Model selection: Cross-validation* [Jupyter Notebook]. Neuromatch Academy: Computational Neuroscience.

Fiveash, A., & Pammer, K. (2014). Music and language: Do they draw on similar syntactic working memory resources? *Psychology of Music*, *42*(2), 190–209. https://doi.org/10.1177/0305735612463949

Goldstein, E. B. (2011). *Cognitive psychology: Connecting mind, research, and everyday experience* (3rd ed.). Cengage Learning.

Granot, R. Y., & Jacoby, N. (2011). Musically puzzling I: Sensitivity to overall structure in the sonata form? *Musicae Scientiae*, *15*(3), 365–386. https://doi.org/10.1177/1029864911409508

Guo, S., & Koelsch, S. (2016). Effects of veridical expectations on syntax processing in music: Event-related potential evidence. *Scientific Reports*, *6*(1), 19064. https://doi.org/10.1038/srep19064

Howell, D. C. (2016). *Fundamental statistics for the behavioral sciences* (9th ed.). Cengage Learning.

Hubbs, N. (1991). Schenker's organicism. *Theory and Practice*, *16*, 143–162.

James, W. (1950). *The principles of psychology*. Dover.

Jan, S. (2010). Memesatz contra ursatz: Memetic perspectives on the aetiology and evolution of musical structure. *Musicae Scientiae*, *14*(1), 3–50. https://doi.org/10.1177/102986491001400101

Justus, T. C., & Bharucha, J. J. (2001). Modularity in musical processing: The automaticity of harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(4), 1000–1011. https://doi.org/10.1037/0096-1523.27.4.1000

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, *9*(4), 637–671. https://doi.org/10.3758/BF03196323

Karno, M., & Konečni, V. J. (1992). The effects of structural interventions in the first movement of Mozart's Symphony in G Minor K. 550 on aesthetic preference. *Music Perception*, *10*(1), 63–72. https://doi.org/10.2307/40285538

Kivy Organization. (2016). *Kivy* (Version 1.9.1). MIT.

Kostka, S., & Payne, D. (2004). *Tonal harmony: With an introduction to twentieth-century music* (5th ed.). McGraw-Hill.

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, *89*(4), 334–368.

Krumhansl, C. L. (2017). Listening niches across a century of popular music. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00431

Krumhansl, C. L., & Zupnick, J. A. (2013). Cascading reminiscence bumps in popular music. *Psychological Science*, *24*(10), 2057–2068. https://doi.org/10.1177/0956797613486486

Laitz, S. G. (2012). *The complete musician: An integrated approach to tonal theory, analysis, and listening* (3rd ed). Oxford University Press.

Lakens, D., & Caldwell, A. R. (2021). Simulation-based power analysis for factorial analysis of variance designs. *Advances in Methods and Practices in Psychological Science*, *4*(1), 251524592095150. https://doi.org/10.1177/2515245920951503

Lalitte, P., & Bigand, E. (2006). Music in the moment? Revisiting the effect of large scale structures. *Perceptual and Motor Skills*, *103*, 811–828.

Lee, Y., Lu, M., & Ko, H. (2007). Effects of skill training on working memory capacity. *Learning and Instruction*, *17*(3), 336–344. https://doi.org/10.1016/j.learninstruc.2007.02.010

Leman, M. (2000). An auditory model of the role of short-term memory in probe-tone ratings. *Music Perception: An Interdisciplinary Journal*, *17*(4), 481–509. https://doi.org/10.2307/40285830

Leman, M., Lesaffre, M., & Tanghe, K. (2001). Introduction to the IPEM toolbox for perception-based music analysis. *Mikropolyphonie-The Online Contemporary Music Journal*, *7*.

Lerdahl, F. (1988). Tonal pitch space. *Music Perception: An Interdisciplinary Journal*, *5*(3), 315–349.

Lerdahl, F. (1992). Cognitive constraints on compositional systems. *Contemporary Music Review*, *6*(2), 97–121. https://doi.org/10.1080/07494469200640161

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

Lerdahl, F., & Jackendoff, R. S. (1996). *A generative theory of tonal music*. MIT Press.

Lerdahl, F., & Krumhansl, C. L. (2007). Modeling tonal tension. *Music Perception: An Interdisciplinary Journal*, *24*(4), 329–366. https://doi.org/10.1525/mp.2007.24.4.329

Lewandowsky, S., Duncan, M., & Brown, G. D. A. (2004). Time does not cause forgetting in short-term serial recall. *Psychonomic Bulletin & Review*, *11*(5), 771–790. https://doi.org/10.3758/BF03196705

Limesurvey GmbH./LimeSurvey. (2020). *An open source survey tool/limesurvey gmbh*. Hamburg, Germany.

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

Marmel, F., Tillmann, B., & Delbé, C. (2010). Priming in melody perception: Tracking down the strength of cognitive expectations. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(4), 1016–1028. https://doi.org/10.1037/a0018735

McAdams, S. (1989). Psychological constraints on form-bearing dimensions in music. *Contemporary Music Review*, *4*(1), 181–198. https://doi.org/10.1080/07494468900640281

Murre, J. M. J., & Dros, J. (2015). Replication and analysis of Ebbinghaus' forgetting curve (D. R. Chialvo, Ed.). *PLOS ONE*, *10*(7). https://doi.org/10.1371/journal.pone.0120644

MuseScore Project. (2015). *Musescore* (Version 2.0).

Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, *18*(3), 251–269. https://doi.org/10.3758/BF03213879

Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology*, *53*(1), 53–81. https://doi.org/10.1146/annurev.psych.53.100901.135131

Neath, I., & Surprenant, A. M. (2003). *Human memory: An introduction to research, data, and theory* (2nd ed.). Thomson/Wadsworth.

Nees, M. A., Corrini, E., Leong, P., & Harris, J. (2017). Maintenance of memory for melodies: Articulation or attentional refreshing? *Psychonomic Bulletin & Review*, *24*(6), 1964–1970. https://doi.org/10.3758/s13423-017-1269-9

Pankhurst, T. (2008). *SchenkerGUIDE: A brief handbook and website for schenkerian analysis*. Routledge.

Pechmann, T., & Mohr, G. (1992). Interference in memory for tonal pitch: Implications for a working-memory model. *Memory & Cognition*, *20*(3), 314–320. https://doi.org/10.3758/BF03199668

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Piston, W. (1941). *Harmony*. W. W. Norton.

Redick, T. S., & Engle, R. W. (2011). Rapid communication: Integrating working memory capacity and context-processing views of cognitive control. *Quarterly Journal of Experimental Psychology*, *64*(6), 1048–1055. https://doi.org/10.1080/17470218.2011.577226

Rogers, W. L., & Bregman, A. S. (1993). An experimental evaluation of three theories of auditory stream segregation. *Perception & Psychophysics*, *53*(2), 179–189. https://doi.org/10.3758/BF03211728

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

Sacks, O. (2007). *Musicophilia: Tales of music and the brain*. Vintage Canada.

Salamé, P., & Baddeley, A. (1989). Effects of background music on phonological short-term memory. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *41*(1-A), 107–122. https://doi.org/10.1080/14640748908402355

Scheirer, E. D. (1996). Bregman's chimerae: Music perception as auditory scene analysis. In B. Pennycook and E. Costa-Giomi (Eds.), *Proceedings of the 4th International Conference for Music Perception and Cognition*, 317–322.

Schenker, H. (1954). *Harmony*. University of Chicago Press.

Schulze, K., & Koelsch, S. (2012). Working memory for speech and music. *Annals of the New York Academy of Sciences*, *1252*(1), 229–236. https://doi.org/10.1111/j.1749-6632.2012.06447.x

Shepard, R. N. (1964). Circularity in judgments of relative pitch. *The Journal of the Acoustical Society of America*, *36*(12), 2346–2353. https://doi.org/10.1121/1.1919362

Snyder, B. (2000). *Music and memory: An introduction*. MIT Press.

Spyra, J., Stodolak, M., & Woolhouse, M. (2021). Events versus time in the perception of nonadjacent key relationships. *Musicae Scientiae*, *25*(2), 212–225. https://doi.org/10.1177/1029864919867463

Spyra, J., & Woolhouse, M. (2021). Influence of surface features on the perception of nonadjacent musical phrases [Manuscript submitted for publication].

Stravinsky, I. (1970). *Poetics of music in the form of six lessons*. Harvard University Press.

Tan, S., & Spackman, M. P. (2005). Listeners' judgments of the musical unity of structurally altered and intact musical compositions. *Psychology of Music*, *33*(2), 133–153. https://doi.org/10.1177/0305735605050648

Temperley, D. (2007). *Music and probability*. MIT Press.

Tillmann, B., & Bigand, E. (1996). Does formal musical structure affect perception of musical expressiveness? *Psychology of Music*, *24*(1), 3–17. https://doi.org/10.1177/0305735696241002

Tillmann, B., & Bigand, E. (2004). The relative importance of local and global structures in music perception. *The Journal of Aesthetics and Art Criticism*, *62*(2), 211–222. https://doi.org/10.1111/j.1540-594X.2004.00153.x

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*(1), 104–132. https://doi.org/10.1037/0033-295X.114.1.104

Woolhouse, M., Cross, I., & Horton, T. (2016). Perception of nonadjacent tonic-key relationships. *Psychology of Music*, *44*(4), 802–815. https://doi.org/10.1177/0305735615593409