

Modeling Melodic Dictation

David John Baker

2018-09-25

Contents

1	Significance of the Study	5
1.1	Claims about need to join the worlds of theory and pedagogy	6
1.2	Chapter Overview	6
2	Theoretical Background and Rationale	9
2.1	What is melodic dictation? and Why?	9
2.2	Individual Factors	13
2.3	Musical Factors	20
2.4	Modeling and Polymorphism of Ability	23
2.5	Conclusions	23
3	History of Aural Skills	25
3.1	Thesis: Show that aural skills always has practical end, efficacy of representation of musical pitch	26
3.2	Quotes from Schumann	26
3.3	Carl Seashore thinking in music	26
3.4	Points from Karpinski on pedagogy	26
3.5	Points from Royal Paper on pedagogy	26
3.6	Solmization System	26
3.7	Really this is all question of efficacy of mental representation of musical pitch	26
4	Individual Differences	27
4.1	Why care about cognitive abilities	28
4.2	Have established that cognitive abilities contribute to musical task (for journal article language repeat)	28
4.3	Remind the nature of a musical dictation type task (hear, loop, executive decision)	28
4.4	WMC has been misused in music education, theory, pedagogy, aural literature and deserves attention	28
4.5	Know WMC plays a role, sense, pertain, execute, should be able to pick up in experiment close to MD	28
4.6	Gold-MSI melodic Memory and beat perception test	28
4.7	IF we accept these DVs, THEN we should be able to predict them with self reports and measures of WMC and gf	28
4.8	Do this with hierarchical LVM ala Elliott paper	28
4.9	Overview of Experiment (cross sectional design)	28
5	Computation Chapter	29
5.1	Humans like patterns and are very good at picking them up	30
5.2	Pre-Musical Corpora	30
5.3	Musical Corpora	30
5.4	So What?	31

6	Hello, Corpus	33
6.1	Brief review of Chapter 4 on corpus (Language to reflect journal submission)	34
6.2	Note problem with using corpus is making corpus	34
6.3	Solem duty to encode and report on corpus	34
6.4	The Corpus	34
6.5	Descriptive Stats of Corpus	34
7	Experiments	35
7.1	Rationale	35
7.2	Experiments	37
7.3	Computational Cognitive Model Model (If time permits) [Whole article in itself]	37
8	Reference Log	43
8.1	To Incorporate	43
8.2	Chapter 3	44

Chapter 1

Significance of the Study

All students pursuing a Bachelor's degree in Music from universities accredited by the National Association of Schools of Music must learn to take melodic dictation (Nat, 2018, Section VIII.6.B.2.A). Melodic dictation is a cognitively demanding process that requires students to listen to a melody, retain it in memory, and then use their knowledge of Western musical notation in order to recreate the mental image of the melody on paper in a limited time frame. As of 2018 there are 647 Schools of Music belonging to National Association of Schools of Music (NASM) CITE WEBSITE, meaning that hundreds of students every year will be expected to learn this challenging task as part of their Aural Skills education. The logic being that as one improves in their ability to take melodic dictation, this practice of critical and active listening develops as a means to improve one's ability to "think in music" and thus become a more competent musician. While learning Aural Skills has been a hallmark of being educated within the Western conservatory tradition, the rationale behind both the how and why of aural skills is often thought of as being esoteric. Throughout the past century, people have disagreed on exactly how one does go about learning a melody with different areas of research each attacking the problem from a different angle.

Despite its ubiquity in curricula within School of Music settings, research on topics pertain to how aural skills are acquired is limited at best. [Citations here about the constant calls butler, klondoski, pembrook] The fields of music theory and cognitive psychology are best positioned to make progress on this question, but often the skills required to be well versed in either of these subjects are disparate, published in other journals, and the research with overlap is scarce. This problem is not new and there have been repeated attempts to bridge the gap between practitioners of aural skills and people in cognitive psychology CITES. Literature from music theory has established conceptual frameworks regarding aural skills Karpinski (2000) and the relevant cognitive psychology literature has explored factors that might contribute to melodic perception (SCHMUKLER SYNER 2016 2016), and there exists applied literature from the world of music education (CITES).

However, despite these siloed areas of research, we as music researchers do not have an a concrete understanding of exactly what contributes to HOW individuals learn melodies (HALPERNBARLETT2010). This is peculiar since "how does one learn a melody" seems to be one of the fundamental questions to the fields of music theory, music psychology, as well as music education. Given this lack of understanding, it becomes even more peculiar that this lack of convergence of evidence is then unable to provide a solid baseline as to what student in their aural skills classrooms can be expected to do. (Also something about we should really know this if we are going to grade people on this ability). While no single dissertation can solve any problem completely, this dissertation aims to fill the gap in the literature between aural skills practitioners (theorists and educators) and music psychologists in order to reach conclusion that can be applied systematically in pedagogical contexts. In order to do this I draw both literatures (music and science) in order to demonstrate how tools from both cognitive psychology as well as computational musicology can help move both fields forward. Some line here about if we really want to understand what is happening we need to know about causal factors going on here and have experimental manipulation and things like making models of the whole thing or talk about what Judea Pearl thinks about the ability to do some sort of causal modeling

with diagrams. Great to rely on some sort of anecdotal evidence, but if we are going to put things on the line with our education then we need to be able to make some sort of falsifiable claims about what we are doing. Can only do that through the lens of science.

1.1 Claims about need to join the worlds of theory and pedagogy

- (Butler, 1997)
- (Klonoski, 2000) - perceptual hierarchy, not enough info from aural skills training
- (Karpinski, 2000) - “There is indeed a gap between the disciples of music cognition and aural skills training”, GK says that one of his goals is to bridge that gap, and he does.

1.2 Chapter Overview

In this first chapter, I introduce the process of melodic dictation and discuss factors that would presumably could play a role in taking melodic dictation. The chapter introduces both a theoretical background and rationale for using method from both computational musicology and cognitive psychology in order to answer questions about how individuals learn melodies. I argue that tools for understanding this best because as we currently understand it, I see us operating in a Kuhnian normal science where much can be learned by just using the tools in front of us. This chapter will clearly outline the factors hypothesized to contribute to an individual’s ability to learn melodies, incorporating both individual and musical parameters. The chapter ends with a discussion some of the philosophical/theoretical problems with attempting to measure things like this (is it just a party trick?) and establishes that I will be taking a more polymorphic view of musicianship in order to answer this question.

The second chapter of my dissertation focuses on the history and current state of aural skills pedagogy.

Tracing back its origins to the practical need to teach musical skills back with Guido d’Arezzo, I compare and contrast the different methodological approaches that have been used, along with their goals.

The third chapter discusses previous work that examines individual factors thought to contribute to one’s ability to perform an aural skills task, and it will discuss results from an experiment contributing to a discussion of how individual differences could contribute to how a person learns melodies.

Turning away from individual differences and focusing on musical features, in the fourth chapter I plan to discuss how music researchers can use tools from computational musicology as predictive features of melodies. Inspired by work from computational linguistics and information theory, recent work in computational musicology has developed software capable of abstracting features thought to be important to learning melodies, such as note density and ‘tonalness’ (Müllensiefen, 2009). Talk a bit about how this has been also looked at before in the music education community.

While these features have been used in large scale, exploratory studies, work in this chapter will discuss how these features could be used in controlled, experimental studies as a stand-in for the intuition many music pedagogues have when determining difficulty of a melody in a classroom setting.

In my fifth chapter, I introduce a novel corpus of over 600 digitized melodies encoded in a queryable format. This dataset will also serve as a valuable resource for future researchers in music, psychology, and the digital humanities. This chapter begins with a discussion of the history of corpus studies, noting their origin outside of music, their current state in music, and their limitations. This chapter, encapsulating the encoding process, the sampling criteria, and the situation of corpus methodologies within the broader research area, will go over summary data and also talk about how it could be used to generate hypotheses for future experiments (n-gram stuff based on patterns) .

Lastly, in the final chapter, I will synthesize the previous research in a series of melodic dictation experiments. Stimuli for the experiments are selected based on the abstracted features of the melodies and are manipulated

as independent variables based on the previous theoretical literature. I then model responses from the experiments using both individual factors and musical features in order to predict how well an individual performs in behavioral tasks similar to some of my previously published research (Baker & Müllensiefen, 2017). Here I also note important caveats in scoring melodic dictation, referencing some other of my own work on using metrics, such as edit distance (Baker & Shanahan, 2018), to discuss similarities between the correct answer and an individual's attempts at dictation. Results from the final chapter will be discussed with reference to how findings are applicable to pedagoges in aural skills settings. Recommendations will be made building on current conceptual frameworks (Karpinski, 2000).

Chapter 2

Theoretical Background and Rationale

2.1 What is melodic dictation? and Why?

Melodic dictation is the process in which an individual hears a melody, retains it in memory, and then uses their knowledge of Western musical notation to recreate the mental image of it on paper in a limited time frame. For many, becoming proficient at this task is at the core of developing one's aural skills (Karpinski, 1990). For over a century, music pedagogues have highly valued melodic dictation¹ which is evident given that most aural skills texts that have sections devoted to honing one's listening skills have material on melodic dictation (Karpinski, 2000). Additionally, any school accredited by the National Association of Schools of Music in North America requires learning this skill be part of any accredited curricula as listed under Section VIII.6.B.2.A.

2.1.1 Small Tirade, could put this at end?

Yet despite this tradition and ubiquity, the rationales as to *why* it is important for students to learn this ability often comes from some sort of appeal to tradition or underwhelming anecdotal evidence that time spent learning to take melodic dictation results in increases in near transfer abilities after an individual acquires a certain degree of proficiency. Rationales given for why students should learn melodic dictation has even been described by Gary Karpinski as being based on “comparatively vague aphorisms about mental relationships and intelligent listening” (Karpinski, 1990, p.192), with the argument for learning to take melodic dictation historically not being well supported with persuasive evidence.

Some researchers have even taken a more skeptical stance and have noted that the rationale for why we teach melodic dictation deserves more critique. For example, Klonoski in writing about aural skills education aptly questions “What specific deficiency is revealed with an incorrect response in melodic dictation settings?” (Klonoski, 2006). Earlier researchers like Potter, in their own publication, have noted that they have been baffled that many musicians do not actually keep up with their melodic dictation abilities after the class ends (Potter, 1990), but presumably go on to have successful and fulfilling musical lives. Additionally, suggesting that people who can hear music and then are unable to write it down, and thus are unable to think *in* music,

¹In his highly influential book *Aural Skills Acquisition: The Development of Listening, Reading, and Performing Skills in College-Level Musicians*, Gary Karpinski (2000) documents this deeply held belief in music pedagogy circles by listing poetic adages from romantic composer Robert Schumann in the mid 19th century through music educator Charles Elliott in the opening of his book, providing concrete examples of the belief that improving one's aural skills, or *ear*, is a highly sought after advanced skill.

seems somewhat exclusionary to the vast amount of musical cultures out there that do not depend on any sort of written notation.

Though despite this skepticism towards the topic, melodic dictation remains at the forefront of many aural skills classrooms. While the act of becoming better at this skill may or may not lead to large increases in far transfer of ability, as a pedagogical tool, teaching students to take melodic dictation brings with it the concepts that are hoped to be instilled in music students long term. While not there has not been extensive research on melodic dictation research in the past few years, in fact (Paney, 2016) notes that since 2000, only four studies were published that directly examined melodic dictation, this skill set sits on the border between literature on learning, melodic perception, memory, and music theory pedagogy and understanding and modeling exactly how it works remains as a untapped watershed of knowledge for the field of music theory, music education, and music perception and is deserving of much more attention.

2.1.2 Describing Melodic Dictation

Much of the foundational work on the topic of melodic dictation comes from the work of Gary Karpinski. Summarized most recently in his *Aural Skills Acquisition* (Karpinski, 2000), though first presented in an earlier article (Karpinski, 1990), Karpinski proposes his own model of melodic dictation that provides a four step framework as how conceptualizes melodic dictation.²

His four steps include:

1. Hearing
2. Short Term Melodic Memory
3. Musical Understanding
4. Notation

and occur as a looping, schematic process that is depicted in Figure 3.1 in his textbook, reproduced below and throughout the chapter Karpinski discusses relevant factors at play at each stage. In the case of the table above, the process happens twice, presumably with a melody that can be extracted in two parts, which according to Karpinski and this prescribed formulas would consist of a melody from twelve to twenty notes long for listeners with few or no chunking abilities. Reviewing the Karpinski four step model will establish the current state of thinking on melodic dictation and thus create a groundwork from which this study will build.

Karpinski's **hearing** stage involves the initial perceptions of the sound at a physical level of the tympanic membrane, as well as the listener's *attention* to the incoming musical information. If the listener is not actively engaging in the task because of factors like "boredom, lack of discipline, test anxiety, attention deficit disorder, or any number of other causes" then any further processes down the model will be detrimentally effected. While Karpinski notes that these types of interferences are normally "beyond the traditional jurisdiction of aural skills instruction", I will later argue that the concept of willful attention when re-conceptualized as working memory capacity, may actually play a large role in the melodic dictation process.

The **short-term melodic memory** Karpinski relies on empirical qualitative evidence from interview with students to show support for two separate memory encoding mechanisms, one for contour, and one for pitch. After citing literature supporting this claim ((Dowling, 1978), (Dewitt and Crowder, 1986)) regarding contour, as well as literature suggesting that in order to memorize chunks of melodic material, it helps to have been enculturated (DEUTSCH 1977, (Oura and Hatano, 1988), Handel 1989, Davies 1978, (Dowling, 1990)) (a field of research that has been explored a lot since this was published (Pearce and Wiggins, 2012)), Karpinski explains how ideas like extractive listening and chunking then take part in this process. Noting that there is probably some sort of capacity limit to music, citing Miller (Miller, 1956) and corroborating musical evidence (Marple n.d), Karpinski details the two strategies from above to show how the dictation process is done. Karpinski's extractive listening refers to only selectively remembering parts of a melody. The chunking bit is an idea that as musical understanding increases, so does musical memory. The idea is that

²In Karpinski 1990 he notes two other models of melodic dictation, one from Rogers where he says there are 2 processes, another from Thomas who says there are 15

if you can identify certain ideas that you hear in the incoming signal, you can more effectively understand them.

- Dowling 1991 (Dowling, 1991)

In order to further explain the idea of **musical understanding** Karpinski says its different ha musical memory. Musical memory for Karpinski is ability to repeat what has been heard, and understanding is about comprehending things like the hierarchical metrical relationships between notes, rhythmic groupings and tonal functions. Within this whole idea is the thing of solimization for both rhythm and pitch (e.g. the *takdimiti* system used with Koda'ly methods) and relative pitch. Idea here is that hte two are then going to be linked to one another and as information comes in that can be musically understood, then it can then be moved from perception to notation more quickly.

The **notation**, and final step of the dictation loop, require the knowledge of Western musical notation to be able to write down what is in one's musical short term memory and use their musical understanding to decode that mental representation into written notation. This last step is ripe for errors as Karpinski notes has been found in other areas of research (Taylor and Pembroke, 1983) and its noted that obviously it's very hard to write something down if you don't have the category for it.

The final parts of the chapter, Karpinski notes that other factors like **tempo**, the **length** and **number of playings**, and the **duration between playings** will also play a role in determining how an individual will perform on a melodic dictation. While this framework can help illuminate this cognitive process and help pedagogues understand how to best help their students, the model as it stands is not detailed enough for explanatory purposes and does not exhaust the many factors that further research since its publication have demonstrated as relevant making a model of melodic dictation.

Again, though Karpinski did not intent to create an exhaustive model of the melodic dictation process, by only creating a verbal model, the model is unprepared to handle a few basic things which are just the next logical research step.

From this point on, need tone to suggest that what Karpinski did is good, but the next step is to get more formal with it and be able to start investigating factors that go into doing melodic dictation beyond just a logical progression of steps. From a resarch and education point of view it would would be good to know how the process works and in order to do that need to look at it as it happens with experimental approaches, and additionally want to try to make a model of it to make sure we include everything.

Title ofis modeling melodic dictation, want to have model that describes it and its component parts, and want to have a model that describes the process.

In this literature review, I use the Karpinski model as a stepping off point by first noting factors that presumably would contribute to this process that ohter people have looked at, organize them in a taxonomy, then review the relevant literature. This is all done in hopes of furthering the fields understanding of melodic dictation and getting a hold of terms so that future research can attempt to be exhaustive in investigating this empirically, exntending the methods used to experimental and computational methodologies.

Where the Karpinski model lacks, using as a scarecrow argument, is the amount of variability at both the individual level, as well as at the musical level.

Primarily, having a single model for melodic dictation assumes that all individuals are likely to engage in this sequential ordering of events. This could in fact be the case, but there is research ANDREW GOLDMAN AND STUFF that suggests that experts process and categorize musical information differently. Additionally, different people will probably have different experiences dictating melodies based on their own past listening experience, a note that Karpinski makes earlier when THAT CITATION FROM UP ABOVE. Not only will there be differences in ability, additionally people differ in their cognitive abilities. The model does not have any flexibility in terms of individual differences.

Secondly, the model also does not seem to treat melodies differently. For example, on page 103, Karpinski suggest that two passes should be adequate for a listen with few to no chunking skills to listen to be able to dictate a melody of twelve to twenty notes. This also would be an area of model improvement, as both the

melodies listed below here are XX notes, within Karpinski’s prescribed range, but intuition might suggest that the melodies are not equally difficult to dictate. The model is agnostic towards differences in melodies.

In order to have a model that will eventually be able to incorporate everything involved in the melodic dictation process, and to be eventually implemented computationally, it is necessary to create a new taxonomy of factors that could contribute to modeling melodic dictation. In the 18 years since the publication of *Aural Skills Acquisition*, there has been a wealth of research that has been done on factors related to melodic dictation. I am convinced that the best way to taxonomize the multitude of factors contributing to this is to create a taxonomy where musical and individual factors are at the highest. By creating a taxonomy of terms it will help organize research and provide a framework to think about how to best operationalize and study factors contributing to modeling melodic dictation.

2.1.2.1 Individual Factors to contribute

From an **individual stand point**, can bifurcate factors broadly into *cognitive factors*, or factors of people that are relatively consistent over people or basically like fixed effects things (use definition of what is a fixed effect); and the other side of this would be things that would have to be dealt with that change with training and exposure. Going to refer to this second set of things as *environmental factors*. Is there a better way to talk about this as nature vs nurture. Additionally need to then mention that there are of course epigenetic factors where both of these parameters might interact with one another. For example, might then imagine that someone with higher cognitive ability, lots of training that was put forward by their parents, as well as tons of musical training, and personality traits that are more likely to learn more (Müllensiefen et al., 2015) might be different in terms of results than someone with lower cognitive abilities, no training, low SES, and a general inclination not to even take music lessons. While obvious, what will eventually be of interest is the degree to which each of these things contributes to the final models. Also gives us a better idea about pedagogy and what not.

2.1.2.2 Musical factors to contribute

In addition to differences at the individual level, there are also **musical level** characteristics. In this category it is also worth taxonomizing the musical characteristics into two categories as well. On one hand we have the **structural aspects** of the melody itself. These are aspects of the melody that would remain invariant when written down on a score. Reading from left to right, would be things like range, key, time signature, intervals, amount of notes, contour of the melody, tonal properties, standardized note density. Then the other side would be musical features that I am going to deem as **experimental features** of the melody. These are aspects of the structure of the melody that you can then warp within the context of a melodic dictation such as tempo, which then reflects note density, timbral qualities, how many times the melody is played, the space between hearings. This is not a categorical divide, while I put something like range as key of the melody as structural, you could imagine that you could have the same interval invariant structure of a melody, perhaps Twinkle, Twinkle little star beginning on C2 notated in bass clef, but then imagine the the same “melody” being played two and a half octaves up on F#4, and transposed to minor and played quicker leading to a phenomenologically similar experience, but not the same. I taxonomize them early on, but again note that a model of this should be able to withstand the multitude of patterns that exist.

2.1.2.3 Make a Model of them

Given all of these factors that then go into the melodic dictation process, the remainder of this chapter will detail previous research that has gone into each of these factors. Talking about each one will provide rationale for why it should be further investigated if we are to better understand melodic dictation. Beginning with *cognitive factors* two levels, then go on to talk about *musical features*. After discussing both and their two subcomponents as I have taxonomized them, offer a brief discussion on how it’s bad to think about these as just latent abilities. Instead talk about thinking about modeling melodic dictation, in terms of ACTUAL

MODELING, as polymorphic conceptualization of aural skills. This is important because how we talk about and model things reflects our values and often it will get in the way of stuff.

This is important because the end goal of this research is to be able to model this from an experimental and computational level.

From an empirical standpoint, both the task as well as the process of melodic dictation as depicted by Karpinski resemble something that can be turned into an experiment, as well as a computational model. It's basically an experiment where we have in person level predictors and melodies with different item level difficulty. This dissertation seeks to explore the degree to which methodologies from cognitive psychology and computational musicology are able to further this literature and take the next logical step in terms of understanding aural skills.

2.2 Individual Factors

2.2.1 Cognitive

Research from the cognitive psychology literature suggests that individuals differ in their perpetual and cognitive abilities in ways that are not easily influenced by short term-training, are stable across a lifetime, and that these abilities, when investigated on a large scale, predict a wealth of human behavior from longevity, to onset of Alzheimer's disease, to the ability to deal with life stress events, and even the effects of alcohol consumption (Ritchie, 2015; ?). Given the strength and generality of these predictors, it's worth investigating the extent that these abilities might play when investigating modeling melodic dictation. This is also important to understand because if we were to find out that something as simple as general fluid intelligence was a strong predictor in musical tasks that we are grading people on, it would be pretty unfair to do since we are then basically grading people on genetic factors beyond their control, as well as environmental factors shown to have some sort of effect on IQ (Ritchie, 2015). Recently there has been a surge of interest in this area (Nancy Rogers, LVH, Utah Guy, Karinski icmpc, Form at SMT) probably due to the fact that educators are picking up on the fact that cognitive abilities are powerful predictors and need to be understood since they inevitably will play a role in pedagogical settings.

Before diving into a discussion regarding differences in cognitive ability, I should note that sometimes ideas regarding differences in cognitive ability have been hostilely received (citation against people talking about IQ) and for good reasons. Research in this area can and has been taken advantage of to further dangerous ideologies (Bell Curve), but often arguments that assert meaningful differences in cognitive abilities between groups are founded on statistical misunderstandings and have been debunked in other literature (Gould, 1996). With that cleared out of the way, it is very hard to maintain a scientific commitment to the theory of evolution (Darwin, 1859) and not expect variation in all aspects of human behavior, with cognition falling within that umbrella.

Attempting to measure aspects of cognition go back over a century. Even before concepts of intelligence were posited by Charles Spearman and his conception of g (Spearman, 1904), scientists were interested in finding links between an individual's mental abilities and some sort of physical manifestation. This area of research was pretty dark and kind of implicitly was all about validating preconceived ideas that people had on the superiority of peoples, with white dudes always coming out on top. For example, THIS GUY WHO TRIED TO QUANTIFY GREATNESS Also this guy who basically measured people's skull sizes. But of course this stuff is not meaningful at all, especially when the dependent variable in this was so subjective and constrained by a culture where one ruling people (white dudes) had all the power.

I only mention this because this line of thinking was co-opted by the American hereditarian school of IQ (page 187 in Gould), where people that were after the same sort of idea (superiority of white people) basically took ideas of Alfred Binet, one of the first people to begin to systematically investigate cognitive ability in children at request of the French government so that children who were, what today we would call learning disabled, could be identified and given special attention. Binet also took a lot of inspiration from Broca, WHO I TALK ABOUT ABOVE ON SKULL SIZE.

Binet was the initial developer of the idea of an intelligence quotient (divide mental age by chronological age then multiply by 100) and provided one of the first ways to attempt to measure something that was not capable of being manifested in the physical world. Around the same time have people like Cyril Burt and Charles Spearman developing new theories of intelligence based on the reification factor analysis and calling it g that is based on solving problems without prior knowledge.

sentence here about g and the positive manifold

Of course took this aside to talk about that basically Binet and Spearman's ideologies about what can be measured still represent two of the largest schools of thought on ways to measure cognitive ability. On one hand there idea that cognitive abilities are based upon a steady growth of incoming information that someone is able to manipulate once they retrieve from long term memory, the other hand is that in addition to that there is some sort of construct in the mind that differs between people that can singularly reflect their intelligence, often referred to as g .

Basic assumption is that there is something in the brain here where differences will lead to different behaviors. Without detailing entire histories of this line of thought, Binet basically turned into an argument for general crystallized intelligence, or DEFINITINO HERE (citation). The Spearman, Burt strain of thought evolved into the g school, and recent research has suggested that the concept of g basically is statistically equivalent to idea conceptualized as general fluid intelligence (cite that in POT). Both of these constructs are powerful predictors on a large scale, but they don't explain the entire picture.

Another large area in the field of cognitive psychology is the area of working memory capacity. As noted in Cowan 2005, what the term refers depends on the framework that is being used, but generally refers to the amount of information that can be actively held in conscious representation, akin to short term memory, but with some important caveats discussed below. In addition to concepts of intelligence, be it Gf or Gc , the working memory capacity literature has done a lot of the heavy theoretical lifting, and after reviewing it we will have a much better idea of how once you decide about how you conceptualize it, then it's helpful.

Given this brief background on cognitive ability, going to now dive deeper into both working memory capacity, as well as general fluid intelligence (avoid calling it g for good reasons) and review literature where these have been discussed as they relate to research on music perception, thus being related to melodic dictation, a skill that you need to perceive music to be able to do?

2.2.2 Working Memory Capacity

Working memory is one of the most investigated concepts in the cognitive psychology literature. According to Nelson Cowan, the term working memory generally refers to

the relatively small amount of information that one can hold in mind, attend to, or, technically speaking, maintain in a rapidly accesible state at one time. The term working is mean to indicate that mental work requires the use of such informaiton. (p.1) (Cowan, 2005)

The term, like most concepts in science, does not have an exact operant definition, nor does it have a definitive method of measurement. While there is no universally recognized first use of the term, researchers began to postulate that there was some sort of system that mediated incoming sensory information with the world with the information in long term storage using modular models of memory in the mid twentieth century. Summarized in (Cowan, 2005), one of the first modal models of memory was proposed by (?) and later expanded by (?). As seen in FIGURE X, both models here posit incoming information that is then put into some sort of limited capacity store. These modal models were then expanded on by Baddeley and Hitch (Baddeley and Hitch, 1974) in their 1974 chapter with the name *Working Memory*, where they proposed a system with an Central Executive module that was able to carry out active maintenance and rehearsal of information that could be stored in either a phonological store for sounds or a visual sketchpad for images.

Later revisions of their model also incorporated an episodic buffer (?) where the modules were explicitly depicted as being able to interface with long term memory in the rehearsal processes. The model has even been expanded upon by other researchers throughout its lifetime. The most notable and relevant to this

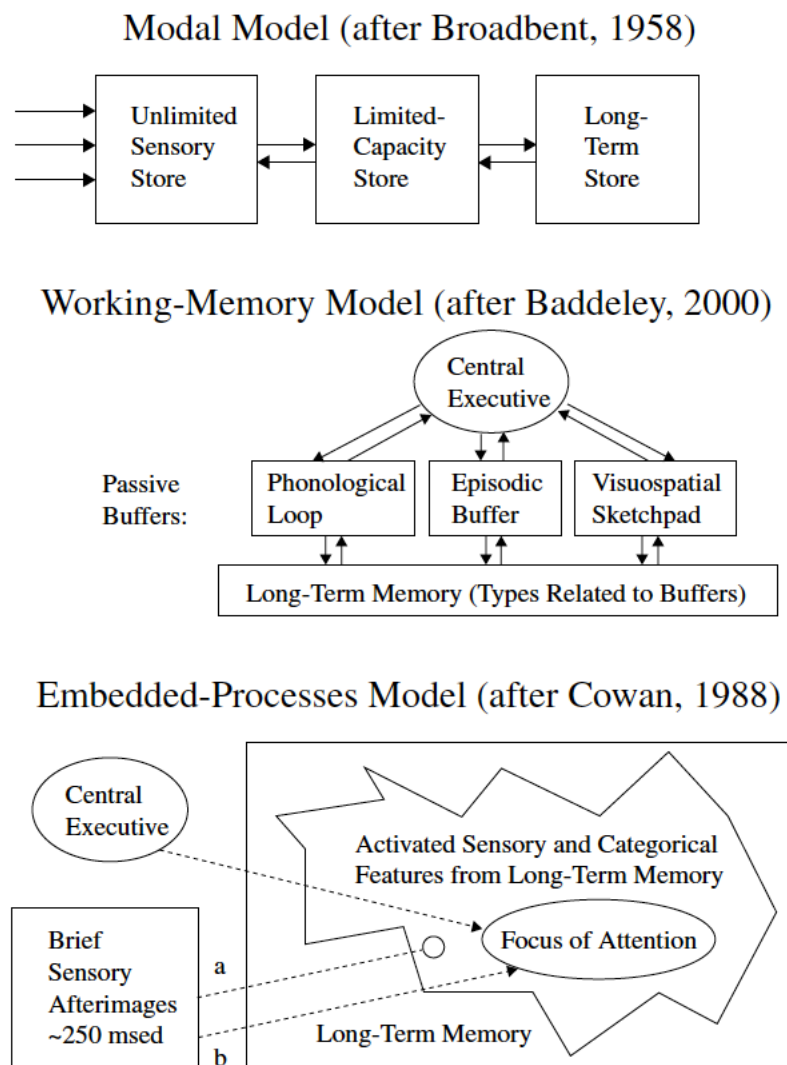


Figure 2.1: Schematics of Models of Working Memory taken from Cowan, 2005

study by is by (Berz, 1995), who postulated adding a musical rehearsal loop to the already established phonological loop and visual spatial sketchpad. While Berz is correct in asserting that the nature of storing and processing musical information is probably different to that of words or pictures and there has been experimental evidence to suggest that they are stored separately (?) and interpreted in favor of multiple loops (?) , it does introduce the theoretical problem of multiple stores which has been addressed by other researchers.

In addressing the problem of explicitly stating which rehearsal loops do and do not exist, Nelson Cowan has proposed other models (Cowan, 2005) such as the Embedded Process Model which do not claim the existence of any domain specific module (e.g. positing a phonological loop, visual spatial sketchpad) but propose an exhaustive model that did away with the problem of asserting specific buffers for new types of information.

In Cowan's own words comparing his model from that of Baddeley:

The aim was to see if the description of the processing structure could be exhaustive, even if not complete, in detail. By analogy, consider two descriptions of a house that has not been explored completely. Perhaps it has only been examined from the outside. Baddeley's (1986) approach to modeling can be compared with hypothesizing that there is a kitchen, a bathroom, two equal-size square bedrooms, and a living room. This is not a bad guess, but it does not rule out the possibility that there actually are extra bedrooms or bathroom, that the bedroom space is apportioned into two rooms very different in size, or that other rooms exist in the house. Cowan's (1988) approach, on the other hand, can be compared with hypothesizing that the house includes food preparing quarters, sleeping quarters, bathroom/toilet quarters, and other living quarters. It is meant to be exhaustive in that nothing was left out, even though it is noncommittal on the details of some of the rooms. p.42

The system is depicted in the bottom tier of TABLE X, and conceptualizes the limited amount of information that is readily available as being in the focus of attention, with activated sensory and categorical features of what is in the focus of attention to be nearly accessible. Moving further from the locus of attention is long term memory, whose content can be turned to by using the central executive to access non immediately available information. In contrast to the modular approaches, Cowan's framework does not require the researchers to specify exactly how and where each the incoming information is being stored which makes it advantageous for studying complex stimuli such as music and melodies.

In addition to having multiple frameworks for studying working memory capacity, there is also the problem of if there are limits to the working memory system, often referred to as the working memory capacity. Most popularized by Miller in his famous (Miller, 1956) article, he suggests out of jest that the number 7 might be worth investigating, which has been used as a point of reference for many researchers since then that have not critically engaged with the article. Miller has even gone on record as noting that using 7 (plus or minus 2) was a rhetorical device used to string together his speech (Miller, 1989). Nevertheless, while probably bogus, this number has been since reduced to about 4 (?) and research around this has been discussed and has been arrived at by using complex span tasks.

When used as predictors in both higher and lower cognitive tasks, measures of working memory capacity predict performance well and additionally tend to be stable across a lifetime (?). Given its predictive strength as well as its direct similarity to tasks of melodic dictation, a in depth look at the literature is warranted.

Clearly an individual's ability to take in sensory information, maintain it in memory, actively carry out other tasks (like noting that melody) are almost identical to tasks of working memory capacity. Although this parallel is striking, typical tasks looking at working memory capacity are different than melodic dictation tasks in a few key ways. The first is that music is always sequential: a melodic dictation task would never require the student to recall the pitches back in scrambled orders. Serial order recall is an important characteristic of the scoring and analyzing of working memory tasks (Conway et al., 2005), musical tones do not appear in random order and are normally in chunks. The use of chunks is pervasive in much of the memory literature, but often is used as more of a heuristic to help explain that information in the

environment is often grouped together. Of the problems with chunking, most are related to music and have relevance to melodic dictation. Below I review the problems with chunking noted by Cowan (Cowan, 2005), and any pertinent music psychology literature.

(Krumhansl, 2001; Meyer, 1956).

1. *Chunks may have a hierarchical organization.* Tonal music has historically been understood to be hierarchical (Krumhansl, 2001; Meyer, 1956; Schenker, 1935) with the study for memory for tones being confounded by some pitches being understood by their relation to structurally more important tones.
2. *The working memory load may be reduced as working memory shifts between levels in hierarchy.* If an individual understands a chunk to be something like a major triad, the load on working memory would be less since it that information could be understood as one chunk.
3. *Chunks may be the endpoints of a continuum of associations.* Pairing a group of tones together that might be functionally anomalous like...
4. *Chunks may include asymmetrical information.* More tonal possibilities are possible from a stable note like tonic or dominant, whereas in a tonal context, a #4 when understood in a functional context would be understood as having less possibilities.
5. *There may be a complex network of associations.* If a set of pitches sounds like a similar set of pitches from LTM, the melody sounds familiar, obviously not just chunk for chunk.
6. *Chunks may increase in size rapidly over time.* Three tones that are seemingly unrelated when incoming like E4, G5, C5 might begin as three different tones, then be fused to be understood as one chunk, a first inversion major triad.
7. *Information in working memory may benefit from rapid storage in long term memory.* Given the amount of patterns that an individual learns and can understand, as soon as something is fused, it could be encoded in LTM, especially if there is some sort of salient feature (immediate recognition of a mode or cadence or whatever)

Important to acknowledge that a lot of this has been done, not do a direct mapping of WMC concepts to it, and proceeding accordingly with the music perception literature that has indicated that we do hear things as statistical learning

- FIND SAFFRAN (Saffran et al., 1999)

This all said, WM literature has done a lot of the theoretical lifting, and WMC is important thing to consider.

Look how Berz ends his article

Important to note here that still area worth looking at because as Berz notes in the last sentence of his article that

Individual differences portrayed in some music aptitude tests may [sic] represent not talent or musical intelligence but ability, reflecting differences in working memory capacity. p. 362

Since it's publication, some studies have looked at this....

2.2.2.1 Papers that suggest WMC plays a role

Of the papers in the music science literature that specifically talk about working memory capacity, each uses different measures of working memory capacity, but all tend to converge on findings that there are some sort of enhanced memory capabilities in individuals with musical training and that working memory capacity, however it is measured, often plays a significant role in musical tasks. Evidence for this comes in a recent meta analyses by Talamini and colleagues (?) who demonstrated via three separate meta-studies that musicians outperform their non-musical counterparts on tasks dealing with long term memory, short term memory, as well as working memory. The authors also noted that the effects were the strongest in working memory tasks where the stimuli were tonal, which again suggests an advantage of exposure and understanding of the hierarchical organization of musical materials. In this meta study and others, it is important to remind that the direction of causality still can not be determined in these types of models.

While it might seem that musical training tends to lead to these increases, it is also equally possible that higher functioning individuals will self select into musical ability or most people will try it out and higher functioning people will be more likely not to quit in the long run.

In terms of musical performance abilities, working memory capacity has also been shown to be a significant predictor. Kopiez and Lee suggested that working memory capacity ought to contribute to sight reading tasks since they found their measures of working memory capacity as measured by a matrix span task to significantly correlated with many of their measures hypothesized to be related to sight reading ability in pianists at lower difficulty grading (??).

Following up on this work on sight reading, Meinz and Hambrick found that working memory capacity as measured by an operation span task, a reading span task, rotation span task, and a matrix span task was able to predict a small amount of variance $R^2 = .074(0.067)$ above and beyond that of deliberate practice alone $R^2 = .451(.441)$ in a sight reading task. More recently, two studies looking at specific sub groups of music have shown working memory capacity to significantly contribute to models of performances on musical tasks. (?) found that although no differences were found between pianists and conductors in measures of working memory capacity as measured via a set of span tasks, conductors showed superior performance in their attention flexibility. Following up on this line of research (Nichols et al., 2018) used the same battery of working memory tasks and found that jazz musicians excelled over their classically trained counterparts in a task which required them to hear notes and reproduce them on the piano. The authors also noted that of their working memory battery, based on standard operation span methods (?), that auditory dilation condition scored surprisingly low and further research on dictation warrants more research. Additionally (?) 2017 found that working memory capacity as measured by a backwards digit span and operation span to be successful predictors in a tapping task requiring sensory motor prediction abilities. Note that each of the tasks mentioned above deal with the the active memory and manipulation of novel musical material and deal with a different type of musical expertise that (e.g., Howe, Davidson, & Sloboda, 1998) argued against the need for WMC at expert level performances.

The growing evidence in this field suggests that the advantage of working memory capacity to be greatest in both musically trained people, dealing with novel information, using tonal materials. Since all three of these factors are related to melodic dictation, it would seem sensible to continue to include these measures in tasks of musical perception and continue Berz's idea that we could just be picking up on individual differences in musical memory.

2.2.2.1.0.1 To improve section on next pass

ADD (?)

2.2.3 General Intelligence

As discussed above gf has a long history, some good some bad. Thing is that it can be really predictive of many things and people def do try incorporate it into music research. And statistically speaking, terms use IQ, but note that IQ and Gf's correlation tend to be statiscally idential (Kovacs and Conway, 2016) and many measures of IQ are done with Ravens as seen below.

One of the big problems with this is establishing effects of causality looking at music. Could be that factors outside of music play a role like personal views of ability (Müllensiefen et al., 2015) , socio economic status, and personality.

2.2.3.1 Papers that suggest GF plays a role

As reviewed in (Schellenberg, 2017), both children and adults who engage in musical activity tend to score higher on general measures of intelligence than their non-musical peers (Gibson, Folley and Park, 2009; Hille et al., 2011; Schellenberg, 2011a; Schellenberg and Mankarious, 2012). with the duration of training

sharing a relationship with the extent of the increases in IQ (Degé, Kubicek and Schwarzer, 2011a; Degé, Wehrum, Stark and Schwarzer, 2015; Corrigan and Schellenberg, 2015; Corrigan, Schellenberg and Misura, 2013; Schellenberg, 2006). Though many of these studies are correlational, they also have made attempts to control for confounding variables like socio-economic status and parental involvement in out of school activities. (Corrigan et al., 2013; Degé et al., 2011a; Schellenberg, 2006, 2011a, 2011b; Schellenberg and Mankarious, 2012). Schellenberg notes the problem of smaller sample sizes in his review (Corrigan and Trainor, 2011; Parbery-Clark et al., 2011; Strait, Parbery-Clark, Hittner and Kraus, 2012) in that they normally do not reach significance. And also references evidence that when professional musicians are matched with non-musicians in the public, also do not seem to see these associations. Schellenberg, 2015 which Schellenberg takes as more evidence that it is actually just higher functioning kids that take music lessons. Additionally, Schellenberg remains skeptical of any sorts of causal factors regarding increases in IQ (e.g., François et al., 2013; Moreno et al., 2009) noting methodological problems like how short exposure times were, or researchers who did not hold pre-existing cognitive abilities constant (Mehr, Schachner, Katz and Spelke, 2013).

- (Corrigan et al., 2013)
- (?)
- (Swaminathan et al., 2017)

2.2.4 Environmental

2.2.5 Long term memory and corpus with implicit

Standing in contrast to factors that individuals do not have a much control over such as the size of their working memory capacity or factors related to their general fluid intelligence, most of the factors we believe contribute to someone's ability to take melodic dictation have to deal with factors related to training and the environment. In fact, one of the tacit assumptions of getting a music degree revolves around the implicitly held belief that with deliberate and attentive practice, that an individual is able to move from novice to expertise in their chosen domain. The idea that time invested results in beneficial returns is probably best exemplified by work produced by ANDERS ERICKSON 1993 that suggests that performance at more elite levels has to do with deliberate practice. Below I review literature that supports this argument, since it's no doubt that someone has to engage in something to be good at it.

2.2.6 Musical Training

Papers that suggest practicing makes you better?

It almost seems redundant to review literature in support of music practice leading to better results.

- List of those papers here

2.2.7 Aural Training

Considering points that I make earlier about polymorphic views of musicality and ability.

- Harrison et al for latent variable approach (Harrison et al., 1994)
- Dictation has not been that well researched (Furby, 2016)
- people try lots of things, not much evidence for peer tutoring (Furby, 2016)

These are things that people have suggested people trying to do : As noted in Furby 2016, researchers in the past have suggested a variety of techniques for improving their abilities in melodic dictation by isolating rhythm and melody [Banton (1995); Bland (1984); Root (1931); WILSON], listening attentively to the melody before writing (Banton, 1995), recognizing patterns (Banton, 1995; Bland, 1984; Root, 1931) and silently vocalizing while dictating (Klonoski, 2006).

2.2.8 Sight Singing

Often described as the other side of the same coin of melodic dictation, sight singing is an area of music pedagogy research that has had sparse attention paid to it given its prevalence in school of music curricula. Recently (Fournier et al., 2017) catalogued and categorized 14 different sub categories into four larger main categories while also providing commentary on some of the current state of aural skills. Of the four large categories, they group them into reading mechanisms, sight singing, readings skills acquisition, and learning support. The authors note a line of research that has documented that university students are often unprepared to sight-read single lines of music (Asmus, 2004; Davidson, Scripp & Welsh, 1988; Fournier, 2015; Thompson, 2004; Vujović & Bogunović, 2012) even though it is, like dictation, thought of as a means for deeper musical understanding. (DeBellis, 2005; Karpinski, 2000; Ottman, 1956; Rogers, 2004; Scripp, 1995; Scripp & Davidson, 1994) The authors of Fournier et. al also note that sight-reading has been an active area of research due to the often reported relationship that performance on sight reading often predicts several studies have shown links between academic success in sight-singing and predictors such as entrance tests (Harrison, 1987, 1990, 1991; Ottman, 1956; Rodeheaver, 1972; Schleuter, 1983), academic ability (Chadwick, 1933; Harrison, 1990, 1991; Harrison, Asmus, & Serpe, 1994; Rodeheaver, 1972), and musical experience (Brown, 2001; Dean, 1937; Furby, 2008; Harrison, 1990, 1991; Harrison et al., 1994; Thostenson, 1967).

Learning to a fluid and competent sight reader presumably bootstraps learning other musical skills, for the same reasons aural skills are taught in schools. Though since it's not directly related and now starting to move away from things looking at dictation directly, time to turn attention to the other half of the taxonomy, the musical factors.

2.3 Musical Factors

Moving to the other half of the taxonomy on figure X, the other big moving part of any study that attempts to look at melodic dictation would be the effects of the melody. Intuitively, everyone has a belief that some melodies are more difficult than others and when stated in the null, with all melodies being equally difficult to dictate, obviously rings true on intuition. According to the way that I was taxonomizing it, the musical factors split into two separate, but not completely split categories. One aspect of this would be the **structural** aspects of a melody. These aspects would be related to discrete pitch and time components of the melody and then any sort of parameter that would result from a certain combination of pitches over time such as contour all which are range invariant. On the other hand would be what I have termed the **experimental aspects** of a melody. The experimental aspects of a melody would be related to aspects of a melody that could be changed that would still make the melody be the same to a reasonable degree like changing the tempo, key, timbre, range, or mode, or even elaborate ornamentation of the melody. Of course there is not a strict delineation between these two categories since one could imagine enough of the experimental parameters modulated to have the melody be categorically different. Questions of musical similarity have been addressed CAMPAPERBOULIS and are beyond the scope of this study.

2.3.1 Structural

- (Schenker, 1935)

The notion that there is something special about the structural aspects of music can be traced back to Leonard Meyer's *Emotion and Meaning in Music* (Meyer, 1956) and has spawned a lot of research since then. Probably one of Meyer's most prolific students was Eugene Narmour

- Eugene Narmour (Narmour, 1990, 1992)

Whose work was expanded upon by David Huron (Huron, 2006) and work using expectational frameworks in melodies was also carried on with work by Marcus Pearce (Pearce and Wiggins, 2012) and Elizabeth Hellmuth Margulis (Margulis, 2005)

- Also schellenberg 1997 (Schellenberg, 1997)

These are general models about melodic perception, though there is a history of looking at melodic dictation specifically. Much of this comes from the world of music education, though not exclusively.

Finally turning to studies specifically looking at melodic dictation looking specifically at musical structure, the first study to examine it extensively was (Ortmann, 1933). Ortmann used a series of twenty five note melodies in order to examine the effects of repetition, pitch direction, conjunct-disjunct motion (contour), interval size, order, and chord structure, all of which he deems to be the *determinants* of an individual's ability to dictate melodic material. This work was extended by * (Taylor and Pembroke, 1983) Extensin of Ortmann and looked at more musical skills, shows ways to suggest scoring simple melodies

Although the literature is generally sparse compared to other areas of music cognition, literature exploring the effects of structural characteristics on memory does exist. Long found that length, tonal structure, contour, and individual traits all contribute and also found that structure and tonalness have significant, albeit small predictive powers in modeling LONG 1977. Problem with long is that they eliminated people who were bad singers for the example. This in some ways might be problematic in that it is at odds both with intuition of representation not being perfect mapping to singing (Pfordresher and Brown, 2007) and that just because someone is bad at singing does not make them bad at representation and the task.

Other researchers have also put forward other parameters like tempo (Hofstetter, 1981) tonality (Dowling, 1978) (Long, 1977) (Pembroke, 1986) (Oura and Hatano, 1988) interval motion (Ortmann, 1933; Pembroke, 1986) length of melody (Long, 1977; Pembroke, 1986) number of presentations (Hofstetter, 1981) [(Pembroke, 1986)] context of presentation (Schellenberg and Moore, 1985) listener experience (Long, 1977; Oura and Hatano, 1988) (Schellenberg and Moore, 1985; Taylor and Pembroke, 1983) familiarity of style (Schellenberg and Moore, 1985)

- (Oura, 1991) Used a sample of eight people to suggest that when taking melodic dictation, individuals use a system of pattern matching that interfaces with their long term memory in order to complete dictation tasks. While this paper does not bring with it a ton of great evidence, the idea is explored in detail in Chapter 6 where I mix up the pattern matching idea of Pearce with Cowan's Embedded process model of for my own model.

More recently

- Gillespie 2001 - HOW DO PEOPLE SCORE MELODIES (Gillespie, 2001)
- Norris 2003 achievement and melodic dictation and sight singing CLEARLY A RELATIONSHIP (Norris, 2003)
- (Pembroke and Riggins, 1990) -
- (Paney and Buonviri, 2014) interviewed HS teachers on how they teach it for AP
- DOOLE DEUTSCH 2010 – people with AP better (Dooley and Deutsch, 2010)
- (Pembroke, 1986) - discussin here wher used tonality, melodic length, and motion, interesting also restricted melodies to ones you can sing. Found length, tonality and motion to be significant at 13 3 1 variance. Though seems like with this there is a lot of chance for type I error and IS THERE EVEN ENOUGH INFO HERE IN THE PAPER TO RECREATE IT ALL. Also claims that people can hear about 10–16 notes, tho DJB would add this is obviously dependent on what i refer to as musical experimental factors and would not make sense to carry this logic forward without more constraints.

2.3.1.1 Recent papers

More recently Nathan Buonviri has published a string of papers where he also has explored melodic dictation as the main fo

Buonviri (2014) found that... Paney and Buonviri (2014) interviewed high school teachers on methods that they used to teach melodic dictation Buonviri and Paney (2015) Buonviri (2015a) Buonviri (2015b)

- (Paney, 2016) found it be beneficial to direct student’s attention and guide them through melodic dictation exercises.
- Buonviri (2017)

Summerize and say that lots of literature has looked at this, but all a bit primitive (change that wording)

2.3.1.2 Caveat about dilberate practice and talent

Also should mention here that clearlry not a perfect one dimensional model where effort equals results. Simple thought expeiment that hard work does not always equal sucess in that there are sucessfull people who have not worked hard and very hard workers who do not experience sucess in what they do. Additionally people need to keep up the skills they have, insert hilarious quote here from potter not believing that people didnt keep up their melodic dictation skills and show that they are compartmentalized (Potter, 1990) This of course does not nullify any findings on deliberate practice, but goes to show that the whole thing is kind of complicated.

2.3.1.3 Recent Computational Musicology Work papers and findings

As noted above, in following studies XYZ, people have looked at musical features before me. Since then there have been a lot further strides in the world of computational musicology.

Since then there has been big moves in the world of computational musicology.

- Huron with kern (Huron, 1994)
- music21 (Cuthbert and Ariza, 2010)

With this new encoding and higher computer power comes ability to do more complicated work. Meyer would have loved it.

While I reserve a longer discussion on the history of computational musicology for the fourth chapter, relevant to this study is the different ways that is possible to talk abstracted features of melodies. Abstracted features of melodies largely be conceptualized as being **static** or **dynamic**. In the above papers, tended to use more simplisitic methods of figuring out parameters, now able to expand using both approaches.

2.3.1.4 Static Views of Computational Features/ FANTASTIC

A static conceptualization of a melody takes a summary feature of it. Helps quantify something that might be intuitive about a melody or piece of encoded music. Do stuff like compare distribution of notes to profiles generated for key finding algorithms (krumhansl, Albrect and shan) Or also has measures of note density. Only problem with this is that not known at this point to what degree the summary features can be mapped on to human behavior.

Quintessential example of this is Feature Analysi... (?). FANTASTIC is software that is capable summerizing musical material at for monophonic melodies. In additon to computing 37 features such as contour variation, tonalnesss, note density, note length, and measures inspired by computational linguistics (THAT BOOK OF GERAINT), FANTASTIC also calculates m-types (melodic-rhythmic motives) that are based on the frequency distributions of melodic segements found genres of music. This is inspired by fact that repetition is key structure of music (Huron, 2006)

Work using the FANTASTIC toolbox has been sucessful in predicting court case decisions (?), predicting chart successes of songs on the Beatles’ *Revolver* (Kopiez and Mullensiefen, 2011), memory for old and new melodies in signal detection experiments (?), memory for earworms (??), memorability of pop music hook (?). In experimental studies, FANTASTIC has also been used to determine item difficulty (Baker and Müllensiefen, 2017; Harrison et al., 2016) and has even been the basis of the development of a computer assisted platform for studyign memory for melodies (?)

2.3.1.5 Dynamic

The other way of conceptualizing music is by using a dynamic approach. Idea here is to take from information content SHANNON and use unsupervised approach.

- Conklin before
- Pearce Model in dissertation 2005
- IDyOM has gone on to be applied in a variety of settings (??)
- Has also been successful in applications to harmony, beating other models PETERS ICMCP PAPER

Advantage to this is that better reflects the real time listening of everything.

2.3.2 Experimental

- Advantage of vocal melodies paper

2.3.3 Brief discussion on Individual and Environment Interactions

Thinking about how all of these parameters contributes makes sense to a certain degree Could imagine the extreme case of each of these either helping or hindering someone's ability to think. Additionally could be the point that things interact with each other. And on top of that we have the idea about levels of Explanation - SEE THAT WITH WIGGINS - SEE THAT WITH HOW COWAN 2005 ends his book

2.4 Modeling and Polymorphism of Ability

Important here to note that unlike Harrison 1994, not really appropriate to conceptualize this as latent ability. Very easy in statistical terms to talk about latent variables and be able to talk about lots of variance predicted. But problem with any sort of model like that or model of musical sophistication is that it has same fate as g.

QUOTES HERE FROM MS paper on latent variables

Musicianship as a concept of something to measure when talking about this one little task with relative lack of concrete evidence for its transfer, need to return back to a process model of melodic dictation.

2.4.1 Polymorphic, component process makes you think about things in models

Going to say here that LVs are bad and that need to adopt a more polymorphic view of musicianship as advocated by

- (Levitin, 2012; Peretz and Coltheart, 2003)

2.5 Conclusions

In this chapter I have first described what is melodic dictation, Karpinski's verbal model of it, noted what the things were that were missing from this model as a stepping off point, then went on to suggest a taxonomy of these based on what already has done. Suggest that there are both individual as well as musical features that are at play here. Individual features can be either cognitive, sort of fixed, or environmental, having to do with training. Musical features can be either structural, or experimental like how you play the melody. None of these are necessarily hard and fast divisions, and certainly there are going to be some interactions between any and all of the levels. Clearly a very great question and given how much we look at it, want

to know more about this complicated network of processes. Given this laundry list of factors, now going to explore the individual parameters in Chapter 3, the musical chapters in 4 and 5, as well as how they would come together in chapter 6. Also will then go into more detail about many of the papers that I mentioned. Really there should be a chapter 7 where after thinking about all of this I put together a computational model that succeeds at incorporating a polymorphic view of musical training by making a process model inspired by Karpinki that is able to be both a research and pedagogical stepping off point for future research on modeling melodic dictation.

Following with this next logical step would be to explore the individual side of the the taxonomy. Transition to working memory capacity chapter...

Chapter 3

History of Aural Skills

3.1 Thesis: Show that aural skills always has practical end, efficacy of representation of musical pitch

3.1.1 for i in star aural people do

3.1.2 Who

3.1.3 Where

3.1.4 When

3.1.5 What

3.1.6 How (approach and goals)

3.1.7 Why

3.1.8 Guido d'Arezzo

3.1.9 Walerant (via Calvisius)

3.1.10 Banchieri

3.1.11 Cerratto

3.1.12 Penna

3.1.13 Zarlino

3.2 Quotes from Schumann

3.3 Carl Seashore thinking in music

3.4 Points from Karpinski on pedagogy

3.5 Points from Royal Paper on pedagogy

Chapter 4

Individual Differences

4.1 Why care about cognitive abilities

4.1.1 General intelligence and WMC

4.1.2 Defining of terms

4.2 Have established that cognitive abilities contribute to musical task (for journal article language repeat)

4.2.1 General Fluid Intelligence, WMC, Training as uni of polymorphic

4.3 Remind the nature of a musical dictation type task (hear, loop, executive decision)

4.3.1 This is WMC task, gf has problems (Although high level link with gf, problematic, WMC models at level of process of md)

4.3.1.1 Berz 1994 noticed it first

4.3.1.2 Williamson Baddely Hitch suggest maybe musical loop

4.3.1.3 Even Cowan labs wonder how different (Li Cowan Saults)

4.4 WMC has been misused in music education, theory, pedagogy, aural literature and deserves attention

4.4.1 Problems with chunking

4.4.1.1 Mistake with Miller 1956, he did not mean 7 items

4.4.1.2 Broadbent 1956 more of why its more like 3-4

4.4.2 Problems with using capacity limit literature

4.4.2.1 See Cowan 2005 page 80

4.4.2.2 Musical order is always serial effects

Chapter 5

Computation Chapter

OTHER PEOPLE WHO HAVE DONE THIS

Folk music

- Bartok 1936?
- Bartok and Lord 1951
- Lomax 1977
- Steinbeck 1982
- Jesser 1992
- Sagrillo 1999
- GET AND READ PAT SAVAGE ARTICLE

Popular Music

- Moor 2006
- Kramarz 2006
- Furnes 2006
- Riedemann ????

Computational Musicology

- Eerola et al 2007 and 2007
- McCay 2005
- Huron 2006
- Frieler 2008
- JAZZOMAT PROJECT OUTPUT

5.1 Humans like patterns and are very good at picking them up

5.1.1 We learn things implicitly

5.1.2 We can represent that implicit knowledge with a corpus

5.2 Pre-Musical Corpora

5.2.1 Information Theory

5.2.2 Computational Linguistics as front runner

5.3 Musical Corpora

5.3.1 History of Musical Corpora

5.3.1.1 Fun old computational music papers

5.3.1.2 Corpora that are often used

5.3.1.3 Static vs Dynamic models of feature abstraction (daniel slides?)

5.3.2 FANTASTIC

5.3.2.1 static

5.3.2.2 ML approach gets it right

5.3.2.3 simple to understand

5.3.2.4 Can abstract features be perceived?

5.3.2.4.0.1 Note density

5.3.2.4.0.2 Contour variation

5.3.2.4.0.3 Tonalness

5.3.2.4.0.4 weird computational measures

5.3.3 IDyOM as representation of musical materials

5.3.3.1 n-gram models

5.3.3.2 mirrors human behavior

5.3.3.2.0.1 melody

5.3.3.2.0.2 harmony

5.4 So What?

5.4.0.1 Other research (Chapt 3) suggest need to move beyond cognitive measures

5.4.0.2 Can operationalize item level items contextually with a corpus

5.4.0.3 IF features are real, they should effect dictation (Chater 6)

5.4.0.4 Not only important for one off, but then would be incorporated into computational learning models (Chapter 6)

5.4.0.5 We need new materials

Chapter 6

Hello, Corpus

6.1 Brief review of Chapter 4 on corpus (Language to reflect journal submission)

6.1.1 Corpus outside of music

6.1.2 Corpus in Music

6.1.3 The point is that it implicitly represents humand knowledge

6.1.4 IDyOM 1

6.1.5 IDyOM 2

6.1.6 IDyOM 3

6.1.7 Huron suggestions that starts of melodies relate to mental rotaiton

6.1.8 Other Huron claims

6.2 Note problem with using corpus is making corpus

6.2.1 Many are used on Essen

6.2.2 Brinkman says Essen Sucks

6.2.3 If going to make generlizable claims, need to always have new data

6.3 Solem duty to encode and report on corpus

6.3.1 Justin London Article on what makes it into a corpusu

6.3.2 Though I just encoded the whole thing because in my heart of hearts I'm a Bayesian

6.4 The Corpus

6.4.1 History of Sight Singign books

Chapter 7

Experiments

7.1 Rationale

7.1.1 Have done all this and have not actually talked about dictation yet

7.1.2 Clearly many factors contribute to this whole thing and need to be taken into a model

7.1.3 Dictation is basically a within subjects design Experiment

7.1.3.1 Get very ecological and dirty and run it

- Paney 2016 had 30 second timing

7.1.4 Factors

7.1.4.1 Cognitive

7.1.4.1.1 WMC

7.1.4.1.2 GF

7.1.4.2 Training

7.1.4.2.1 Goldsmiths MSI

7.1.4.3 Musical

7.1.4.3.1 FANTASTIC

7.1.4.3.2 IDyOM

7.1.4.4 Investigate melodies with this context and set scoring

7.1.4.5 Mirror design to see if effects of melody are there

7.2 Experiments

7.2.1 Experiment I

7.2.1.1 Participants

7.2.1.2 Procedure

7.2.1.3 Materials

7.2.1.4 Scoring

7.2.1.5 Results

7.2.1.6 Modeling

7.2.1.7 Discussion

7.2.2 Experiment II

7.2.2.1 Participants (New)

7.2.2.2 Procedure (Same)

7.2.2.3 Materials (Swapped but controlled)

7.2.2.4 Scoring (Same)

7.2.2.5 Results

7.2.2.6 Modeling (same)

7.2.3 General Discussion

7.2.3.1 What happened

7.2.3.2 Assumption of all of this is that many things are happening linearly in combination with each other

7.2.3.3 Additionally the mixed effects framework works better with more data?

7.2.3.4 Also how we score it is going to mess with the DVs

7.2.4 Really what is needed is Computational Model

7.3 Computational Cognitive Model Model (If time permits) [Whole article in itself]

So thinking all about this, first reviewing the literature and then trying to run experiments to figure it all out, got kind of frustrated because even with a mixed effects model, all it is saying is that as note density goes UP so does difficult, and it interacts with tonality. Def setting myself up for a “I could have told you that” moment from my music theory colleagues. Also figured since I am doing a dissertation called modeling melodic dictation, would be good to actually make an explanatory model of it.

Hopefully by this point, have a good idea of the factors that might contribute, my task here was to then try and operationalize everything and make a computer do it.

Actually really inspired by this Lewandowsky book where they had a quip about how wit the Baddeley model it really has like 150 permutations of it and some famous example of someone saying that men have more one night stands than women, but in a heterosexual population that is mathematically impossible. So it only seemed like it made sense to see if I could write one out.

So this is how I think it would work, currently in the process of coding it.

So essentially it’s a bayesian inspired model that compares a target melody, with a corpus of melodies or the prior, that are meant to represent all “known” musical material in someone’s musical understanding. I got the idea from reading about Cowan’s Embedded process model which does not posit the working memory structure as very distinct from that of everything in long term representation, but rather uses the central executive idea as the limited window of attention that can then be focused on things either in LTM or near categorical features of what is in the purview of attention.

Will first go through it on a high level, then make another pass and talk about how everything is getting operationalised. So imagine that we have a prior corpus of all the melodies that are covered for sight singing and dictation in the first semester of aural skills (in this case, melodies X–XX). If we let the corpus M represent all prior knowledge, where not only is the melody represented as a string of intervals, but we also get every n gram permutation of of the melody based on the idea that the more frequent the exposure of an n-gram, the more likely it would be to be understood.

This is the first big parameter M.

We could imagine and visualize this as a grid of distributions of all possible n-grams and the strings that make them up seen here GGLOT2. Now although any string of notes is possible, in order to mirror the limited capacity of WMC, there needs to be some sort of threshold that is set that would mirror the limits of the window of WMC.

This threshold is the next big parameter, T.

Currently conceptualizing this as being reflective of maximum of information content threshold as calculated by IDyOM.

Idea here is that as you hear more notes, more IC, the more your fixed capacity bin of melody is going to be filled up. By making it based on the IC of a melody string represented in the corpus, patterns that are more frequent are going to be easier to then dictate, which aligns with intuitions on melodic dictation.

Also need to define a the explicit, implicit threshold. Explicitly known n-grams should have some sort of fixed amount that if it succeeds that, you would know it. Like karpinski “knowing” phase, would also assume that every interval class would also be “known” and if not, that is OK too, will lead to people not being able to transcribe it.

We’ve now defined both the individual knowledge parameters and the threshold T of working memory capacity. These consist of WHAT DO YOU CALL THIS.

With these individual parameters established, we can calculate difficulty of the melody by running what I am going to be determining as the ‘transcribe’ function, which is separate from the individual parameters.

We then can introduce the target melody , t which represents the string to be dictated.

Idea here is that this melody is compared to the corpus, M, and you get the IC of each of the n-grams from the prior corpus.

Given the threshold, take the biggest n-gram that fits below the threshold to be put in the buffer. If you had a huge threshold and could take IC of whole melody or you KNEW it, would be reflective here. Could also introduce a function here that just guesses at about the threshold. Once this is filled up, let's say 4 notes, the notes are then set to the transcription buffer.

At this point index the n-gram is cross checked against the prior knowledge and look for matches of it that happen above a certain "known" threshold. If it is known, it gets "notated" and you re-enter the melody. If not, in this case the 4 gram would get recursively truncated until found a string, even 2 gram where it is known. If it gets to a 2 gram and it's not explicitly known, basically because you don't know that interval, what Karpinski rails against as atomistic hearing. Also good in that people with bigger chunking will do better.

If there is not an exact match, move to split fuzzy searches (future versions)

Re-entry function is defined by finding the next part of the melody where either next most common n-gram occurs or where it left off last. This is then to reflect the two different ways of approaching tacking. Seing as some people start at front, others start at end. Could be reflective of primacy and recency effects of memory, OR maybe it's due to fact that more prototypical n-grams in habit starts and ends of melodies....?

This process would then re-occur a few times over until either can't go anymore (intervals are not known) or till completion.

Afterwards have an idea of the path that it took, counts of those, use those as proxies of difficulty.

I think this covers everything in the decision making process and each step of the process can be automated eventually (post doc anyone?!). More importantly, it reflects ecological phenomenological experience of taking melodic dictation.

Aligns with intuitions that People who know more melodies will be better (singer phenomena) People will try for bigger chunks first, then go to smaller ones Atomistic transcription can happen but is inefficient People with higher chunking ability will do better IC is helpful proxy for this in line with IDyOM literature Relative pitch is where it's at for things like this What's deal with AP?

7.3.1 Why?

7.3.1.0.1 Better than verbal models

7.3.1.0.2 Sometimes even mathematically infesable proposed theory

7.3.1.0.3 Beyond Karpinski in that it doesn't just schematize, says exactly when each thing is happening when

7.3.1.0.4 Lends itself to better discussions that don't just rely on personal anecdotes

7.3.1.0.5 Can tweak the parameters

7.3.1.0.6 Can collect different types of data (corpus or experimental) and use the model

7.3.1.0.7 This model suggests that atomism approach is actual just subprocess of larger pattern

7.3.1.1 Theoretical Justification

7.3.1.1.1 Marries literature on LTM and prior knowledge, information theory, WMC, computation, representation

7.3.1.1.2 Also can be implemented in computer

7.3.1.1.3 representation of rhythm too?

7.3.1.1.4 inspired by people like margulis 2005, albrecht and shanahan key finding, want something to contribute

7.3.1.1.5 Really Made me think

7.3.1.2 The Model (note many parameters can be changed in R package)

7.3.1.3 Prior

7.3.1.3.0.1 Corpus of music represented in form of n-grams

7.3.1.3.0.2 IDyOM extracts all possible n-gram permutations as learned corpus

7.3.1.3.1 Music notation fed into processing window where incoming n-gram is matched based on WMC window OR IT maximum

7.3.1.3.1.1 Information builds until approaches critical threshold

7.3.1.3.1.2 Upon maximum, model puts n-gram into focus of attention (Cowan 1988) and note why this is better than Baddely Hitch

7.3.1.3.1.3 Recursive transcribe function looks for LTM matches

Option 1: Pattern Matched and Pattern Transcribed, success?

Option 2: Pattern not matched in full, truncated and use match option again (should be higher probability of match with corpus)

Option 3: Pattern not matched downsize again until at interval level and relying on 2-gram (atomism)

On success of option, reopen gate at nearest long implicit n-gram LTM Match (start or end problem)

7.3.1.3.1.4 Put time constraints on search features

7.3.1.3.1.5 Transcribe process resets with trace image of melody after each dictation

7.3.1.3.1.6 Transcribe process ends when all notes accounted for

7.3. COMPUTATIONAL COGNITIVE MODEL MODEL (IF TIME PERMITS) [WHOLE ARTICLE IN ITSELF]⁴¹

7.3.1.4 Model Output

7.3.1.4.1 Based on learning, times needed to hear it

7.3.1.4.2 Completion percentage

7.3.1.4.3 Rank order of easier to transcribe parts based on learning

7.3.1.5 Model Compared to Data

7.3.1.5.1 With Experimental Data

7.3.2 Future Suggestions for Aural Skills Pedagogues and Research

7.3.2.1 Use model as teaching stepping off point

7.3.2.2 Should move towards LTM pattern matching

7.3.2.3 Reason that people learn how to sight sing is to INCREASE the learning of the implicit corpus

7.3.2.4 Circular process here

7.3.2.5 Is this what it means to then think IN music

7.3.2.6 Really it's to just know the patterns maybe like model where Justin London suggests we get to know patterns and expect them

7.3.2.7 Would also make sense in terms of Leonard Meyer 1956

7.3.2.8 Use WMC in music theory, cognition, education studies

Chapter 8

Reference Log

8.1 To Incorporate

- (Margulis, 2005) – Margulis Model
- (Nichols et al., 2018) – Specialty jazz background helps in tasks, WMC
- (?) – Fix intext
- (Schumann and Klauser, 1860) – Quote about why people should do ear training
- (Smith, 1934) – Quote from K2001 about why people should do ear training
- (Long, 1977) – Musical Characteristics predict memory
- (Taylor and Pembroke, 1983) – Great citation that lots of things change memory, even structural!
- (?) – Long boring talk on STM, LTM
- (Oura, 1991) – Awful experimental design that says people use structural tones
- (Buonviri, 2014) – Call for experimental, suggestions as to what factors might contribute, use of deductive reasoning, qualitative
- (Buonviri, 2015b) – People need to focus right away, not establish, distractors
- (Buonviri, 2015a) – Showing people visual music does not help much.
- (Buonviri, 2017) – Listening helps with other things, no best strategy in terms of writing
- (Buonviri and Paney, 2015) – Literature to say people are bad at teaching melodic dictation and we don't know a lot about it, also interesting stuff about what solfege systems people use
- (Butler, 1997) – Call for music educators to do aural skills research, notes problem with aural skills pedagogy in lack of direction, also nice Nicholas Cook quotes on point of theory
- (Furby, 2016) – music ed study with weird stats, has references to follow up on with advantages of pitch systems and people who recommend things for sight singing
- (Pembroke, 1986) – Effects of melodies, also how people do it. Interesting that they too effect of melodies, but talk about things in terms of notes and not in terms of information content. Thought of have an experiment where the n-grams that are more common are easier to write down. Lots of good charts too.
- (Paney, 2016) – It's not good if you tell people what to do when they are dictating, article has a lot of good review for dictation materials to add to the 'toRead' folder.
- (Fournier et al., 2017) – Good references that people are awful at Aural Skills, Also suggestions that people are not that great at transfer, and some stuff to suggest academic ability is intertwined in all of this. Good reference for when starting to talk about untangling the mess that is aural skills.
- (?, 1995) – Add on a new module to the WMC model of baddel with music, presents some evidence for why this theoretically should be included, but actually takes examples of dictation. A lot of this article felt like things that i was reinventing...not good.
- (?) – Proof some other people are starting to think in terms of pedagogical schemas
- (Klonoski, 2000) – Music cognition needs to talk to aural skills more, also need to unbind theory routine with aural skills and think of things more as in a perceptual learning hierarchy

- (Klonoski, 2006) – great quotes that when people get something wrong with aural skills, what does that even mean, lack of transfer effects, article ends with ways to get better at things
- (Pembroke and Riggins, 1990) – Survey of what people in the late 1980s were doing in terms of aural skills pedagogy
- (?) – addresses why Gary Karpinski thinks we should teach melodic dictation
- (Potter, 1990) – dictation teacher surprised that people don't keep up their dictation skills quote

8.2 Chapter 3

- (Cowan, 2005) – This book will probably serve as cornerstone of chapter in terms of creating relevant literature in addition to EE course readings on WMC. Provides history of WMC models and notes how attention based model as opposed to Baddeley loop might actually be better theoretical model for talking about fact that WMC could just be something related to attention if not that. Provides extensive listing on problems with chunking that are all relevant to music, but then also supports it. Shows that Miller 1956 is a generally bad citation, own author even says that in Miller 1989 (check and add) and says limit is probably about 4 (use Cowan 2001 for citation find that). Lots of good ideas like how music is always serial recall, examples of how to model the process, great discussions on zooming out and categorical nature of music within span of WMC ideas.
- (Ockelford, 2007) – uses case of savant to argue bits of Berz WM Music Model

Bibliography

- (2018). National Association of Schools of Music Handbook. Technical report, National Association of Schools of Music, Reston, Virginia.
- Baddeley, A. D. and Hitch, G. (1974). Working Memory. In *Psychology of Learning and Motivation*, volume 8, pages 47–89. Elsevier.
- Baker, D. J. and Müllensiefen, D. (2017). Perception of Leitmotives in Richard Wagner’s Der Ring des Nibelungen. *Frontiers in Psychology*, 8.
- Banton, L. (1995). The Role of Visual and Auditory Feedback during Sight-Reading of Music. *Psychology of Music*, 23(1):3–16.
- Berz, W. L. (1995). Working Memory in Music: A Theoretical Model. *Music Perception: An Interdisciplinary Journal*, 12(3):353–364.
- Bland, L. (1984). *Sight Singing Through Melodic Analysis: A Guide to the Study of Sight Singing and an Aid to Ear Training Instruction*. Scarecrow Press.
- Buonviri, N. (2015a). Effects of music notation reinforcement on aural memory for melodies. *International Journal of Music Education*, 33(4):442–450.
- Buonviri, N. O. (2014). An Exploration of Undergraduate Music Majors’ Melodic Dictation Strategies. *Update: Applications of Research in Music Education*, 33(1):21–30.
- Buonviri, N. O. (2015b). Effects of a Preparatory Singing Pattern on Melodic Dictation Success. *Journal of Research in Music Education*, 63(1):102–113.
- Buonviri, N. O. (2017). Effects of Two Listening Strategies for Melodic Dictation. *Journal of Research in Music Education*, 65(3):347–359.
- Buonviri, N. O. and Paney, A. S. (2015). Melodic Dictation Instruction. *Journal of Research in Music Education*, 62(2):224–237.
- Butler, D. (1997). Why the Gulf between Music Perception Research and Aural Training? *Bulletin of the Council for Research in Music Education*, (132).
- Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., and Engle, R. W. (2005). Working memory span tasks: A methodological review and user’s guide. *Psychonomic Bulletin & Review*, 12(5):769–786.
- Corrigall, K. A., Schellenberg, E. G., and Misura, N. M. (2013). Music Training, Cognition, and Personality. *Frontiers in Psychology*, 4.
- Cowan, N. (2005). *Working Memory Capacity*. Working memory capacity. Psychology Press, New York, NY, US.
- Cuthbert, M. S. and Ariza, C. (2010). Music21: A Toolkit for Computer-Aided Musicology and Symbolic Music Data. page 7.

- Darwin, C. (1859). *On the Origin of Species*. Routledge.
- Dewitt, L. A. and Crowder, R. G. (1986). Recognition of Novel Melodies after Brief Delays. *Music Perception: An Interdisciplinary Journal*, 3(3):259–274.
- Dooley, K. and Deutsch, D. (2010). Absolute pitch correlates with high performance on musical dictation. *The Journal of the Acoustical Society of America*, 128(2):890–893.
- Dowling, W. J. (1978). Scale and Contour: Two Components of a Theory of Memory for Melodies. *Psychological Review*, 84(4):341–354.
- Dowling, W. J. (1990). Expectancy and attention in melody perception. *Psychomusicology: A Journal of Research in Music Cognition*, 9(2):148–160.
- Dowling, W. J. (1991). Tonal strength and melody recognition after long and short delays. *Perception & Psychophysics*, 50(4):305–313.
- Fournier, G., Moreno Sala, M. T., Dubé, F., and O’Neill, S. (2017). Cognitive strategies in sight-singing: The development of an inventory for aural skills pedagogy. *Psychology of Music*, page 030573561774514.
- Furby, V. J. (2016). The Effects of Peer Tutoring on the Aural Skills Performance of Undergraduate Music Majors. *Update: Applications of Research in Music Education*, 34(3):33–39.
- Gillespie, J. L. (2001). Melodic Dictation Scoring Methods: An Exploratory Study. *Journal for Music Theory Pedagogy*, 15.
- Gould, S. J. (1996). *The Mismeasure of Man*. WW Norton & Company.
- Harrison, C. S., Asmus, E. P., and Serpe, R. T. (1994). Effects of Musical Aptitude, Academic Ability, Music Experience, and Motivation on Aural Skills. *Journal of Research in Music Education*, 42(2):131.
- Harrison, P. M., Musil, J. J., and Müllensiefen, D. (2016). Modelling Melodic Discrimination Tests: Descriptive and Explanatory Approaches. *Journal of New Music Research*, 45(3):265–280.
- Hofstetter, F. T. (1981). Computer-Baased Recognition of Perceptual Patterns and Learning Styles in Rhythmic Dictation Exercises. *Journal of Research in Music Education*, 29(4):265–277.
- Huron, D. (1994). *The Humdrum Toolkit: Reference Manual*.
- Huron, D. (2006). *Sweet Anticipation*. MIT Press.
- Karpinski, G. (1990). A Model for Music Perception and Its Implications in Melodic Dictation. *Journal of Music Theory Pedagogy*, 4(1):191–229.
- Karpinski, G. S. (2000). *Aural Skills Acquisition: The Development of Listening, Reading, and Performing Skills in College-Level Musicians*. Oxford University Press.
- Klonoski, E. (2000). A Perceptual Learning Hierarchy: An Imperative for Aural Skills Pedagogy. *College Music Symposium*, 4:168–169.
- Klonoski, E. (2006). Improving Dictation as an Aural-Skills Instructional Tool. page 6.
- Kopiez, R. and Müllensiefen, D. (2011). Auf der Suche nach den “Popularitätsfaktoren” in den Song-Melodien des Beatles-Albums Revolver: Eine computergestützte Feature-Analyse [In search of features explaining the popularity of the tunes from the Beatles album Revolver: A computer-assisted feature analysis]. *Musik und Popularität. Beiträge zu einer Kulturgeschichte zwischen*, pages 207–225.
- Kovacs, K. and Conway, A. R. A. (2016). Process Overlap Theory: A Unified Account of the General Factor of Intelligence. *Psychological Inquiry*, 27(3):151–177.
- Krumhansl, C. (2001). *Cognitive Foundations of Musical Pitch*. Oxford University Press.

- Levitin, D. J. (2012). What Does It Mean to Be Musical? *Neuron*, 73(4):633–637.
- Long, P. A. (1977). Relationships Between Pitch Memory in Short Melodies and Selected Factors. *Journal of Research in Music Education*, 25(4):272–282.
- Margulis, E. H. (2005). A Model of Melodic Expectation. *Music Perception: An Interdisciplinary Journal*, 22(4):663–714.
- Meyer, L. (1956). *Emotion and Meaning in Music*. University of Chicago Press, Chicago.
- Miller, G. A. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information.
- Miller, G. A. (1989). *A History of Psychology in Autobiography*, volume VIII. Stanford University Press, Stanford, CA, 1. gardner edition.
- Müllensiefen, D., Harrison, P., Caprini, F., and Fancourt, A. (2015). Investigating the importance of self-theories of intelligence and musicality for students' academic and musical achievement. *Frontiers in Psychology*, 6.
- Narmour, E. (1990). *The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model*. University of Chicago Press, Chicago.
- Narmour, E. (1992). *The Analysis and Cognition of Melodic Complexity: The Implication-Realization Model*. University of Chicago Press.
- Nichols, B. E., Wöllner, C., and Halpern, A. R. (2018). Score one for jazz: Working memory in jazz and classical musicians. *Psychomusicology: Music, Mind, and Brain*, 28(2):101–107.
- Norris, C. (2003). The Relationship Between Sight Singing Achievement and Melodic Dictation Achievement. *Contributions to Music Education*, 30(1):39–53.
- Ockelford, A. (2007). A Music Module in Working Memory? Evidence from the Performance of a Prodigious Musical Savant. *Musicae Scientiae*, 11(2_suppl):5–36.
- Ortmann, O. (1933). Some tonal determinants of melodic memory. *Journal of Educational Psychology*, 24(6):454–467.
- Oura, Y. (1991). Constructing a Representation of a Melody: Transforming Melodic Segments into Reduced Pitch Patterns Operated on by Modifiers. *Music Perception: An Interdisciplinary Journal*, 9(2):251–265.
- Oura, Y. and Hatano, G. (1988). Memory for Melodies among Subjects Differing in Age and Experience in Music. *Psychology of Music*, 16:91–109.
- Paney, A. S. (2016). The effect of directing attention on melodic dictation testing. *Psychology of Music*, 44(1):15–24.
- Paney, A. S. and Buonviri, N. O. (2014). Teaching Melodic Dictation in Advanced Placement Music Theory. *Journal of Research in Music Education*, 61(4):396–414.
- Pearce, M. T. and Wiggins, G. A. (2012). Auditory Expectation: The Information Dynamics of Music Perception and Cognition. *Topics in Cognitive Science*, 4(4):625–652.
- Pembrook, R. G. (1986). Interference of the Transcription Process and Other Selected Variables on Perception and Memory during Melodic Dictation. *Journal of Research in Music Education*, 34(4):238.
- Pembrook, R. G. and Riggins, H. L. (1990). "Send Help!": Aural Skills Instruction in U.S. Colleges and Universities. *Journal of Music Theory Pedagogy*, 4(1):230–241.
- Peretz, I. and Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, 6(7):688–691.

- Pfordresher, P. Q. and Brown, S. (2007). Poor-Pitch Singing in the Absence of "Tone Deafness". *Music Perception: An Interdisciplinary Journal*, 25(2):95–115.
- Potter, G. (1990). Identifying Successful Dictation Strategies. *Journal of Music Theory Pedagogy*, 4(1):63–72.
- Ritchie, S. (2015). *Intelligence: All That Matters*. All that matters. Hodder & Stoughton.
- Root, F. (1931). *Methodical Sight-Singing: Lessons in Vocal Culture*. Bryn Mawr, Pennsylvania: Theodore Presser Co.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., and Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70(1):27–52.
- Schellenberg, E. G. (1997). Simplifying the Implication-Realization Model of Melodic Expectancy. *Music Perception: An Interdisciplinary Journal*, 14(3):295–318.
- Schellenberg, E. G. (2017). Music and Nonmusical Abilities. page 17.
- Schellenberg, S. and Moore, R. S. (1985). The Effect of Tonal-Rhythmic Context on Short-Term Memory of Rhythmic and Melodic Sequences. *Bulletin of the Council for Research in Music Education*, (85):207–217.
- Schenker, H. (1935). *Der Freie Satz*, volume III. Universal Edition.
- Schumann, R. and Klauser, K. (1860). *Musikalische Haus-Und Lebens-Regeln*. J. Schuberth & Comp.
- Smith, M. (1934). Solfège: An essential in musicianship. *Music Supervisors' Journal*, 20(5):16–20.
- Spearman, C. (1904). "General Intelligence," Objectively Determined and Measured. page 93.
- Swaminathan, S., Schellenberg, E. G., and Khalil, S. (2017). Revisiting the association between music lessons and intelligence: Training effects or music aptitude? *Intelligence*, 62:119–124.
- Taylor, J. A. and Pembroke, R. G. (1983). Strategies in memory for short melodies: An extension of Otto Ortmann's 1933 study. *Psychomusicology: A Journal of Research in Music Cognition*, 3(1):16–35.