Modeling Melodic Dictation

SMT40-Prospectus

David John Baker

October 26th, 2018

Motivation

Melodic dictation is the process in which an individual hears a melody, retains it in memory, then uses their knowledge of Western musical notation to recreate the mental image of the melody on paper in a limited time frame. For many, becoming proficient at this task is at the core of developing one's aural skills. This is evident from the fact that most aural skills textbooks with content devoted to honing one's listening abilities have sections devoted to learning how to take melodic dictation (Karpinski 2000). Additionally, any school accredited by the National Association of Schools of Music in North America requires students to learn this skill ("National Association of Schools of Music Handbook" 2018, sec. VIII.6.B.2.A). Yet, despite this ubiquity in the pedagogy of melodic dictation, exactly *how* this process works is not understood: Music theorists do not have an explanatory theory of melodic dictation.

The lack of knowledge regarding melodic dictation is alarming given the degree music theorists are engaged with the teaching and assessing of this ability. As a community, a more systematic understanding of *how* people learn melodies is not only important from an pedagogical point of view, but understanding how people learn and perceive melodies is at the locus of research central to music theory, music education, as well as music cognition. While there have been repeated calls throughout the past few decades to synthesize these disparate literatures (Butler 1997; Karpinski 2000; Klonoski 2000), the literature is sparse in relation to how frequent melodic dictation appears as part of our curricula.[1] Reviewing the current state of research on melodic dictation highlights the need for the music theory community to better understand melodic dictation and the literature that surrounds it.

Much of the fundamental work on melodic dictation was synthesized via the work of Gary Karpinski. Originally appearing in an article from 1990, and then later the focus of of the third chapter of *Aural Skills Acqusition*, Karpinski proposes a four step model that describes an idealized process of melodic dictation (Karpinski 2000, 1990). The four steps include:

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

and are conceptualized as a looping process that is done over each chunk of musical material that the listener focuses on via a process of extractive listening. His flowchart of the process is reproduced below in Figure 1.

As a pedagogical tool, Karpinski's model distills a complicated and almost esoteric process into a manageable system that benefits both students as well as aural skills pedagogues. Karpinski's model describes the process of melodic dictation but his model makes no claims as to *how* the process happens. Though not the original intention of the model, this four step model lacks robustness in that it is agnostic to both differences at the individual level, as well as for differences in melodic material.

For example, Karpinski suggests an example for discussion based on his idealized process and claims that listeners with "few to no chunking skills" should be able to dictate a melody of twelve to twenty notes long with two passes of his flowchart. While this provides an approximate rule of thumb as to what can be expected of students, these suggestions are generated from a fixed system and are not flexible to individuals

[1] Paney (2016) notes that since 2000 only four studies explicitly examining melodic dictation have been published in music pedagogy journals.

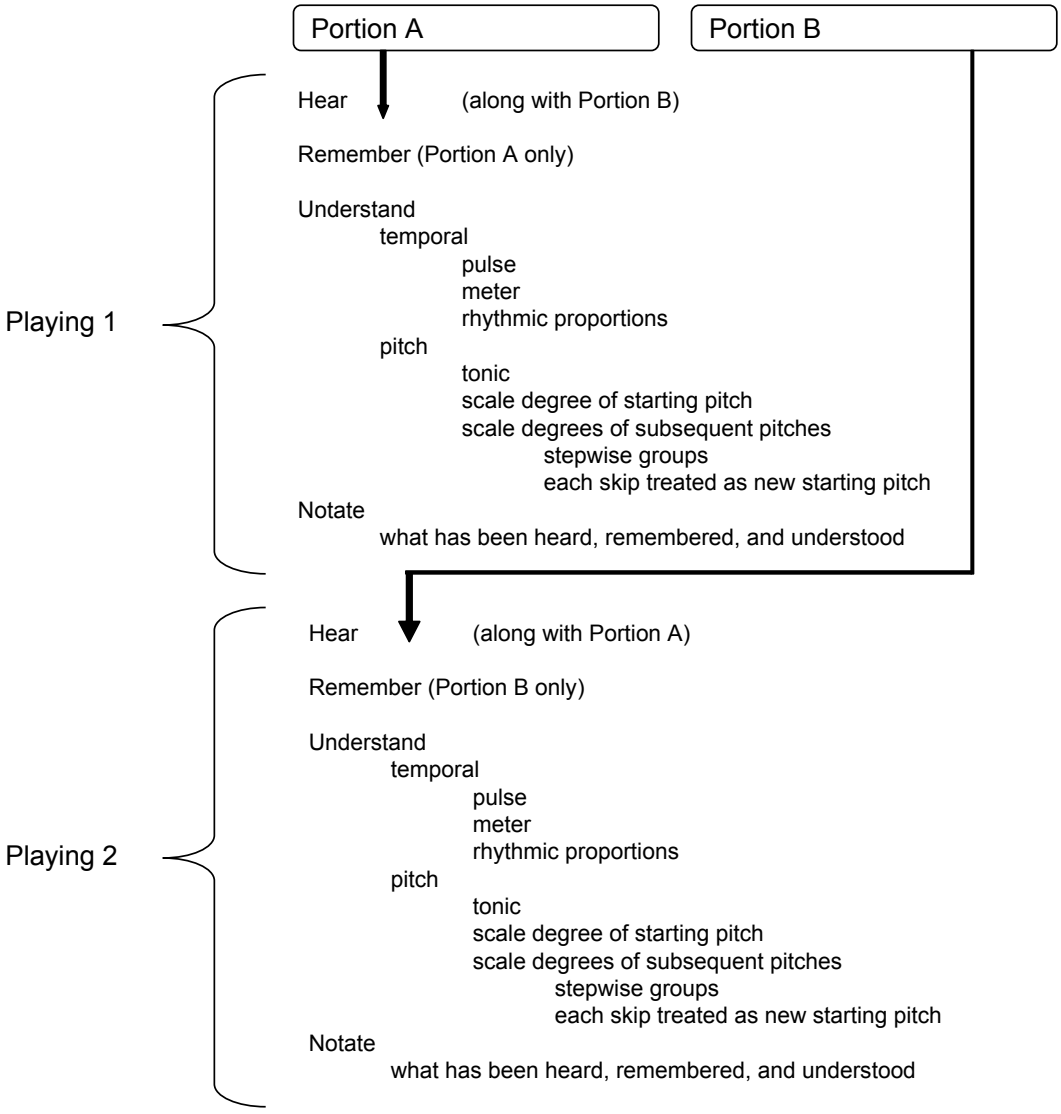

Figure 1: Karpinski Idealized Flow Chart of Melodic Dictation

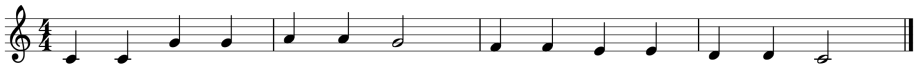

Figure 2: Figure 2: Melodies of Equal Length: Tonal

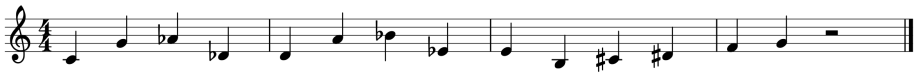

Figure 3: Figure 3: Melodies of Equal Length: Atonal

with different experience levels.[2] To give an example, both melodies from Figure 2 and 3 are 14 notes in length, yet the strategies and thus perceived difficulty of dictating each melody would be presumably different for individuals with different musical training backgrounds. Not only will an individual's prior experience affect this process, but presumably the melodic material will as well.

In this dissertation, I continue the line of research begun with Gary Karpinski's four step model of melodic dictation; I do this by first exploring how both individual experiences and musical material can affect melodic dictation separately, examine how these factors interact, then finally posit an explanatory computational model of melodic dictation.

The past 20 years of research since *Aural Skills Acquisition* have highlighted the importance of taking into account both ability to understand how individual differences play a role in music perception tasks as well as an ability to quantify and operationalize differences in melodies that reflect a theorists' intuitive understanding of melodic material.

In order to organize and then reflect on the vast amount of factors that could contribute to how a person performs in melodic dictation I propose a taxonomy– seen in Figure 4– of parameters with both individual (e.g. cognitive and environmental) and musical (e.g. structural and experimental) parameters from approximately the last two decades that should be further explored when looking at melodic dictation in order to move towards an explanatory theory of melodic dictation.

Using this taxonomy as a guide, I investigate factors thought to contribute to tasks of melodic dictation using a diverse methodological toolbox which borrows techniques ranging from cognitive psychology, to computational musicology, and music theory.

[2]Karpinski does in fact note that as listeners develop more varied skills they can "entertain some significant deviations from such a process," p.103 and gives relevant examples, but these deviations are not formalized which is most likely due to the multiplicity of permutations that may exist.

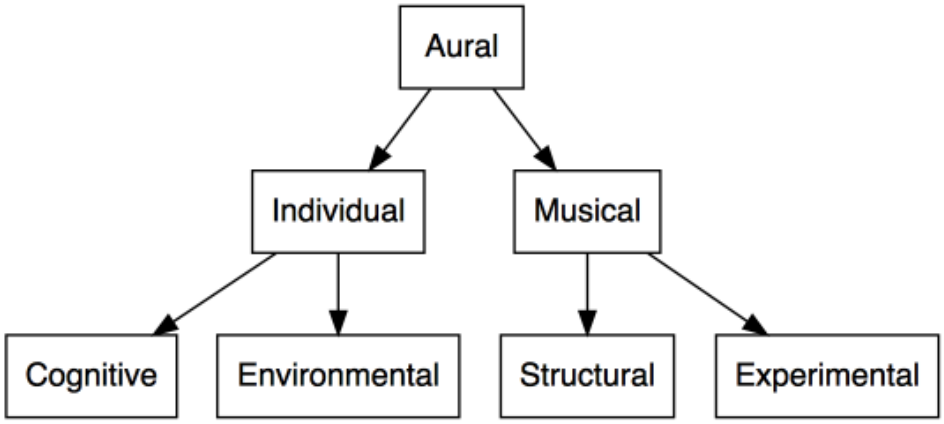

Figure 4: Taxonomy of Factors Relating to Melodic Dictation

Present Research Question

This dissertation takes an interdisciplinary approach to create an explanatory theory of melodic dictation. In order to do this, I synthesize and utilize work from music theory, music education, and music cognition. Since the goal of this research is to improve the music theorists ability to be effective pedagogues, I frame my research question in relation to answering questions as they relate to undergraduate students.[3] Specifically, I write five chapters that respectively investigate the following five questions:

1. To what degree does a student's attention and cognitive ability determine how well they can perform melodic memory tasks?
2. How can computationally abstracted melodic features help determine the difficulty of dictation melody in line with expert intuition?
3. Can patterns from a corpus of sight singing melodies serve as representations of musical structures that students know both implicitly and explicitly?
4. If conceptualized as an experiment, is it possible to predict how a student does on a melodic dictation task when both individual and musical features are accounted for? What aspect is the most predictive of an individual's ability? What problems arise in operationalizing a melodic dictation exercise?
5. Is it possible to posit a computational, explanatory model of melodic dictation that mirrors the phenomenological decision process individuals engage with when performing melodic dictation? If so, how can this be used to help teach this process?

With each of these five questions serving as the basis for each of the five chapters of original content, I next detail what I will use for each of these methodologies in detail.

[3]Although pedagogically focused, the research here also has more domain general applications and relevance.

Outline Research Methods

In order to investigate my first research question, I analyze and interpret data from a large scale experiment conducted over the last year in collaboration with the Louisiana State University's cognitive psychology program. The experiment uses a large sample of (N = 470) students who took participated in a multi part experiment where we took measures of cognitive ability using multiple measures of for both general fluid intelligence and working memory capacity, measures of musical background via subjective self report, and then also an objective test of musical memory via a melodic memory test. Given the complexity of variables at play, I analyze the data using structural equation modeling– a statistical technique developed in order to parse out causal relationships within covariance structures– and find evidence that working memory plays a large role in tasks of musical perception. These results corroborate earlier theoretical claims regarding the need to investigate musical working memory (Berz 1995).

I then introduce a novel corpus in the third chapter that will consist of over 600 sight singing melodies. Using recent advancements in computational musicology, I show how there are statistical norms present in this corpus and suggest that there is a link between using these melodies as an individual's implicit knowledge of a musical structure. In the fourth chapter I demonstrate how computationally extracted features can serve as a quantifiable proxy for a pedagogues' intuition and propose application for future research. Following earlier claims by Meyer (1956) and later computationally driven research by Pearce (2018) I note how information derived from these corpus studies can be incorporated in experiments investigating melodic dictation. This corpus can also be utilized by other researchers who wish to investigate corpus level claims.

Synthesizing the assumptions and findings from the previous three chapters, I combine the work on individual differences and melodic differences in order to operationalize every factor of relevance discussed in earlier chapters in order to conduct an experiment. Here I reflect on the choices needed to be made in this process and talk about implications and limitations of measuring melodic dictation in an experimental setting. Using these experimental methods allows the researcher to determine what happens when the aforementioned musical materials are manipulated in an ecologically valid setting. Findings from this chapter corroborate earlier findings that musical features do in fact play a large role in an individual's ability to perform melodic dictation(Ortmann 1933; Pembrook 1986; Taylor and Pembrook 1983). In detailing the methodologies used in this chapter, I also put forward a complete and free set of software and analysis method so future researchers looking to explore this question on their own are not limited in terms of technical accessibility.

Finally, in the last chapter I put forward a computational model of melodic dictation that explicitly details each step of the melodic dictation process. This Bayesian inspired model takes account of a listener's prior knowledge, explicit understanding of musical material, and then allows a melody to be computationally dictated in order to result in an difficulty score based on previous knowledge, as well as measures thought to be influential like working memory capacity from above. In formalizing this as a computational model, I account for both musical and individual factors relevant to the melodic dictation process, thus providing a new pedagogical and research framework to situate work on melodic dictation.

Significance of Project to Music Theory

Exploring melodic dictation using interdisciplinary approaches allows the music research community to have a more comprehensive understanding of melodic dictation. Firstly, and most importantly and as stated above, there is not an explanatory model of melodic dictation. Discussed in more detail in the attached sample chapter, much of the current work on melodic dictation is supported using descriptive models that only assert what will happen in response to changing one or two aspects of a melodic dictation. Contrary to a descriptive model like Karpinski's Four Step Model, an explanatory model not only posits *what* will happen, but hypothesizes the inner workings of the process and asserts both *how* and *why*. Foreshadowing this included chapter, I combine recent advances in computational musicology via the work of Marcus Pearce's IDyOM model (Pearce 2018)– essentially a computational implementation of theoretical claims made by Meyer (1956)– and Nelson Cowan's Embedded Process model (Cowan 1988, 2010) in order to put forward full, explanatory model of melodic dictation.

Given its importance to the responsibilities of most music theorists, a working theory of how melodic dictation will further our pedagogical understandings. Additionally by putting forward an explanatory theory in the form of a computational model that is informed by experimental evidence, theorists are able to discuss the melodic dictation process with more exact language. Though initially conceptualized as a model for melodic dictation, the model might also have additional predictive power due to its basis in a domain general memory mechanism and could demonstrate how insights from music theory could be used to help explain how the mind works

In terms of novel research findings, this dissertation provides new evidence for some earlier theoretical claims. As mentioned above, work from this dissertation suggests that working memory does in fact play a large role in questions of melodic perception, thus corroborating earlier claims by Berz (1995). This dissertation also puts forward a new corpus of over 600 melodies that can be used by the computational musicology community. Additionally, having written software using open source libraries, new avenues of research can be now open to other researchers interested in setting up experiments investigating melodic dictation with little to no programming experience. Finally, this dissertation posits a new, and explanatory model of melodic dictation based on work from cognitive psychology and computational musicology. In addition to the rationale above, the music theory pedagogy community might be able to more effectively teach melodic dictation. By having a step-by-step model detailing the exact processes used, it will become easier to discuss issues related to melodic dictation like the efficacy of one sight singing system over another, how to best invest time practicing skills related to melodic dictation, and what can be reasonably expected in the assessing of students on this complex ability.

The model also puts forward a series of hypotheses that could be falsified in order to investigate the model's verisimilitude. Among other hypotheses the model predicts:

- Segments of melodies are likely successfully to be dictated relative to the frequency distribution of their prior knowledge.
- Higher working memory span individuals will be able to dictate bigger chunks of melodies, and thus perform better at dictation
- Using an *atomistic*[4] dictation will result not as effective dictations than attempting to identify larger patterns
- Determining the difficulty of melodies of equal length is predictable from the frequency the melody's cumulative n-gram distribution.
- Some *atonal* melodies will be easier to dictate than tonal melodies if they consist of patterns that are more frequent in a listener's prior knowledge
- Higher exposure to sight-singing results in more explicitly learned patterns, thus the ability to identify larger patterns of music

Although many of these hypotheses might seem intuitive to any instructor that has taught aural skills before, work from this dissertation provides a theory as to why each appears to be true. Future research beyond this dissertation will explore further predictions of this work in more detail.

As stated above, given the ubiquity of melodic dictation in Schools of Music and our responsibility as music theory pedagogues to be able to explain what we do, it is important to have a theory of melodic dictation that is explanatory so students can be correctly assessed on this exercise. As a community it is important to know the degree to which abilities can be learned and how much ability derives from pre-existing individual differences so that we as pedagogues can cater to the diversity of students that we teach. This is especially important given current conversations regarding the content what should be considered central to the curricula for an accredited School of Music education. Throughout the dissertation–though not a central thesis– I argue for a more modular, polymorphic conception of musicianship based on my work on melodic dictation. Discussed mostly in the first chapter literature review and then touched on when relevant, I deconstruct concepts like musicianship and musical sophistication by paralleling recent arguments that deconstruct the notion of a general intelligence factor in psychological research. I argue that conceptualizing musicianship as a set of related, though not unified abilities, we as pedagogues can reconceptualize how we teach and think

[4] A term used by Karpinski to describe the process of trying to hear a melody by listening interval by interval

about the processes that enable or musical choices and be more effective teachers. This argument will appear as part of a forthcoming article that I authored in *Musicae Scientiae*.

Current Progress

As of October 26th, three chapters of the dissertation have been completed as drafts. The completed chapters include the literature review, the fifth chapter on experimental method, as well as the attached chapter that describes the computational model. The second chapter currently exists as a conference proceedings paper from the International Conference on Music Perception and Cognition 15 and I am currently in the process of changing the language and expanding on all of the figures and analyses used in the shorter version of the paper. I am still in the process of encoding the melodies and have finished encoding one third of the melodies in the corpus. The chapters on using computational tools as well as describing the corpus have been outlined, but have not been completed.

Research that resulted from earlier work investigating problems with measuring and modeling musical ability has been accepted for publication and is forthcoming in *Musicae Scientiae* and work from the fifth chapter on using experimental methods as a means to understand melodic dictation is set to be published in a chapter for an upcoming *Routledge* book that has been tentatively titled *Understanding Melodic Dictation via Experimental Methods*. Once finished, the computational chapters are set to be submitted to the *International Society for Music Information Retrevial* and results from the experiments are set to be published as an article that includes both sets of experiments for a music perception journal. A review of melodic dictation literature as it relates to Karpinski's work as well as the computational model as it pertains to music theory pedagogy will also be published in a journal where the readership overlaps between music theory, music cognition, and music education. Current progress of the dissertation is available online where all the materials, software, text, and supplemental literature cited in the dissertation can also be accessed in able to facilitate more accessible research for anyone looking to become involved with this area of research.

A completed draft of the dissertation will be sent to the committee in early 2019, thus allotting two and a half months for revisions before the official submission date of March 18th. The dissertation will be defended in the weeks after submission at a time where all members of the committee are able to attend.

Select Bibliography

Berz, William L. 1995. "Working Memory in Music: A Theoretical Model." *Music Perception: An Interdisciplinary Journal* 12 (3): 353–64. https://doi.org/10.2307/40286188.

Butler, David. 1997. "Why the Gulf Between Music Perception Research and Aural Training?" *Bulletin of the Council for Research in Music Education*, no. 132.

Cowan, Nelson. 1988. "Evolving Conceptions of Memory Storage, Selective Attention, and Their Mutual Constraints Within the Human Information-Processing System." *Psychological Bulletin* 104 (2): 163–91.

———. 2010. "The Magical Mystery Four: How Is Working Memory Capacity Limited, and Why?" *Current Directions in Psychological Science* 19 (1): 51–57. https://doi.org/10.1177/0963721409359277.

Karpinski, Gary. 1990. "A Model for Music Perception and Its Implications in Melodic Dictation." *Journal of Music Theory Pedagogy* 4 (1): 191–229.

Karpinski, Gary Steven. 2000. *Aural Skills Acquisition: The Development of Listening, Reading, and Performing Skills in College-Level Musicians*. Oxford University Press.

Klonoski, Edward. 2000. "A Perceptual Learning Hierarchy: An Imperative for Aural Skills Pedagogy." *College Music Symposium* 4: 168–69.

Meyer, Leonard. 1956. *Emotion and Meaning in Music*. Chicago: University of Chicago Press.

"National Association of Schools of Music Handbook." 2018. Reston, Virginia: National Association of Schools of Music.

Ortmann, Otto. 1933. "Some Tonal Determinants of Melodic Memory." *Journal of Educational Psychology* 24 (6): 454–67. https://doi.org/10.1037/h0075218.

Paney, Andrew S. 2016. "The Effect of Directing Attention on Melodic Dictation Testing." *Psychology of Music* 44 (1): 15–24. https://doi.org/10.1177/0305735614547409.

Pearce, Marcus T. 2018. "Statistical Learning and Probabilistic Prediction in Music Cognition: Mechanisms of Stylistic Enculturation: Enculturation: Statistical Learning and Prediction." *Annals of the New York Academy of Sciences* 1423 (1): 378–95. https://doi.org/10.1111/nyas.13654.

Pembrook, Randall 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. https://doi.org/10.2307/3345259.

Taylor, Jack A., and Randall G. Pembrook. 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. https://doi.org/10.1037/h0094258.