STANFORD UNIVERSITY

Correlation analyses of encoded music performance

by

Jeffrey C. Smith

A thesis submitted in partial fulfillment for the degree of Doctor of Philosophy

in the
Center for Computer Research in Music and Acoustics
Stanford Department of Music

November 29, 2013

Bartók recording folk songs from Slovak peasants in 1907.
(Getty Images, De Agostini Picture Library)

"In a book on ethnomusicology, we cannot avoid adding the example of our principal tool: the phonograph."

Curt Sachs

Abstract

Correlation analyses of encoded music performance by a large and diverse international community of amateur performers afford insights into fundamental questions of musical behavior. By observing demographic data associated with a corpus of performance recordings, we can conjecture about cultural, geographical, topographical, sociopolitical, economic and other potential influences, and explore possible 'universals' in musical thought and practice.

A century ago, Béla Bartók visited what were then remote regions to identify and characterize folk music at its sources. Subsequent investigators and collectors, with a growing awareness of the effects of colonialism and with a variety of objectives, sought to describe and categorize music from a broad spectrum of cultures and regions (Agawu). In some cases western music was introduced to an indigenous populace specifically in order to observe and record listener reaction (Sachs 17).

Although studies of this sort are fraught with issues[1], there is a great deal to gain by examining how a particular populace interprets a foreign object. In terms of cultural objects such as a work of music, inferences can be drawn as to what (if anything) is 'universal', as well as how to characterize cultural differences.

As music delivery, through increasingly pervasive mobile devices, becomes more available as well as more interactive, new opportunities arise to study musical practices. Interactive applications merge audio playback with recorded performance, effectively providing users novel musical instrument interfaces that are amenable to mastery by amateurs. These software instruments are limited in their acoustic richness and expressive control. Their available sounds and tunings are also generally biased towards mainstream western musical practice. On the other hand, within a relatively short time and with minimal frustration, users both young and old can learn to perform a wide range of pre-composed works in diverse styles and genres. More importantly, despite (perhaps, because of) these instruments' limitations, two aspects of performance that are especially revealing of competence and style — tempo and agogics — are accurately recorded and efficiently stored.

The underlying premise of this thesis is that, embedded in the encoded music performances by this diverse community of amateur performers is a wealth of information about musical performance practices. Correlation analysis of these performances can

[1]Ultimately, the study of what (if anything) unites the seemingly vast and diverse musical practices amongst humans, as well as how to characterize difference, is hampered by what Agawu calls the 'us/them' dichotomy.

provide insight into musical behaviors as a whole, as well as comparative observations on the cultural, social, economic and geopolitical influences on music performance.

This study attempts to address questions of musical practices from an entirely novel perspective, specifically taking advantage of massively popular game-oriented music performance programs on mobile devices.

Acknowledgements

I owe a debt of gratitude to many of the faculty members at Stanford (and BYU) who helped me learn about music. First and foremost, I must thank my advisor and mentor, Jonathan Berger. I recall meeting with Dr. Berger in the fall of 2005 to discuss whether I might be considered for the PhD program. Despite inherent vagaries, I was admitted to CCRMA by the grace of God. It was Dr. Berger's insight into composition, music analysis, and performance practices that ultimately helped me formulate this thesis. George Barth, who served on my committee and with whom I have studied piano performance and music theory for many years, offered new perspectives on both music and life that transformed my approach to performance and interpretation — another motivation for this thesis on the study of performance practices. To Dr. Robert B. Smith, I will never tire of the Kabalevsky Piano Concerto we did or Prokofiev's 7th Sonata, or of course Bach's Italian Concerto. To Adolph "Ushu" Baller, the slow movement of Beethoven's first piano concerto that we played at your studio in your home studio in Palo Alto is always with me. Thank you for your love of music and your enduring kindness to me as my teacher for eight years. And thank you for all of the German cookies you gave me after every lesson. To Chris Chafe who served as the co-chair on my committee, who introduced me to computer music, and who created an environment at CCRMA that I cherish, "that the free, exploring mind of the individual human is the most valuable thing in the world," (Steinbeck, East of Eden), I thank you Chris. To Giancarlo Aquilanti, who served on my committee, taught me music theory for three quarters and music composition on the side, and wrote some of the most beautiful music I experienced at Stanford; to Craig Sapp whose research showed the potential for the statistical analysis of performances and whose guidance and mentorship helped enable much of the research I undertook; to Kumaran Arul who opened my mind to the possibilities; to Ge Wang who, while living in a cold, dark, secretive room with an iPad chained to a desk, built the products that gathered the data that made this thesis possible; to Greg Waite who taught me how to sing; to Robert Huw Morgan whose many performances in Memorial Church helped me grieve; to Mark Applebaum for his creativity, music, art, insight, and friendship; to Bryan Wolf who chaired my oral committee; and to the rest of the faculty who taught, encouraged, and supported me, I thank you. Special thanks to my colleagues Nick Rudolfsky for building the initial shell to load the piano data, and Michael Wang, Michael Berger, and Rob Hamilton for their help in creating and accessing much of the performance data. And last, thank you to my very first music teacher, Mom.

Contents

List of Figures

List of Tables

Abbreviations

MIDI	**M**usical **I**nstrument **D**igital **I**nterface
AAT	**A**ggregate **A**verage **T**empo
BPM	**B**eats **P**er **M**inute
EsAC	**E**ssen **A**ssociative **C**ode for monophonic music
Pcc	**P**earson **c**orrelation **c**oefficient
A-Pcc	**A**verage **P**earson **c**orrelation **c**oefficient
DAMP	**D**igital **A**rchive of (Mobile) **M**usical **P**erformances

Dedicated to Dad.

Chapter 1

The elusive search for 'universal' musical behaviors

Nagy csukott ajtókat látok

1.1 motivation

An evening in Transylvania

A century ago, Bartók and Kodály ventured into remote Hungarian villages, recording folk music at its source with their primitive phonograph. Bartók persevered in his effort to identify and preserve intrinsic music through the final years of his life, even after leaving Hungary and moving to New York (Stevens). Why?

We know that Bartók criticized several contemporary musicologists' characterizations of "Hungarian" and "Gypsy" music. He claimed they were "reversing the natural order of things, in which practice precedes theory" (Stevens 47). In some ways, Bartók's motives might be attributed to the most basic principles of academic research, and his desire to understand the underlying 'primitives' of musical materials and gestures. Alternatively, or perhaps additionally, he may have desired to capture and preserve intrinsic music before the effects of globalizing technologies and demographic changes rendered isolated cultures vulnerable.[1]

[1] It has been suggested that Bartók and Kodály may also have been championing the unique characteristics of Hungarian music, reflecting the trend towards musical nationalism prevalent in regions outside of Germany and France at the end of the Romantic Era (Taruskin 376-378). Yet we know that Bartók. unlike Kodály, went beyond Hungary into neighboring regions and even North Africa because of a conviction that was at odds with nationalism:

> Bartók reached the conclusion that the richest folk music came from groups whose heritage was impure, groups that had rubbed elbows with other national or racial groups over

Indeed, in the decades that followed, the possibility of documenting cultural practices unperturbed by outside influence became increasingly remote.

> In 1950 Gyula Kertesz and Zoltan Kodály undertook the recording of peasant music in the district around *Mohacs*, discovering that many singers who had been recorded in the 1930's were still accessible. Comparison of the records of the two periods disclosed a pronounced deterioration of the peasant music, even in so short a time. (Stevens)

Shostakovich described an episode of deliberate destruction of intrinsic cultural practices during this same era in the former Soviet Union.

> I hope someone will write down the history of how our great native art was destroyed in the twenties and thirties. It was destroyed forever because it was oral. When they shoot a folk singer or a wandering storyteller, hundreds of great musical works die with him. Works that had never been written down. They die forever, irrevocably, because another singer represents other songs.. . .
>
> Since time immemorial, folk singers have wandered along the roads of the Ukraine. They're called *lirniki* and *banduristy* there. They were almost always blind men — why that is so is another question that I won't go into, but briefly, it's traditional.. . .
>
> And then in the mid thirties the First All-Ukrainian Congress of *Lirniki* and *Banduristy* was announced, and all the folk singers had to gather and discuss what to do in the future. "Life is better, life is merrier," Stalin had said. The blind men believed it. They came to the congress from all over the Ukraine, from tiny, forgotten villages. There were several hundred of them at the congress, they say. It was a living museum, the country's living history. All its songs, all its music and poetry. And they were almost all shot, almost all of those pathetic blind men killed. (Shostakovich)

Sachs cited other factors that accelerated the trend towards the erosion of isolated cultures.

a considerable period of time. His visit to North Africa confirmed this theory, when he found the music of the Arabs, isolated in vast reaches of the Sahara, less highly developed and consequently less interesting than that of the Magyars and the surrounding peoples. The avoidance of foreign influences, he concluded, whether deliberate or not, leads to stagnation; enrichment of folk music results from the absorption of such influences. (Stevens 48)

> In pursuing our goal [the preservation of indigenous music], we are particularly rushed, as the venerable heritage of archaic cultures is threatened with imminent extermination. The melodies of primitive man, an organic, essential part of his spiritual life, fall victim to Christian missionaries, Soviet agents, European colonizers, and American oil drillers. And primitive as well as folk melodies are hopelessly succumbing to a technical age with military service and factory work, with rapid buses, planes, and cars, with phonographs, radios, and television sets. (Sachs 3)

Yet, unlike Sachs, Bartók came to appreciate the 'impurities' in his recordings and transcriptions of folk music. The folk music Bartók collected and studied had a profound impact on his compositions (Taruskin). Examples abound, from his *Evening in Transylvania* from *Ten Easy Piano Pieces*, to the *Scherzo alla bulgarese* from the Fifth String Quartet. Unlike Bartók the ethnomusicologist, Bartók the composer adapted folk elements, abstracting and generalizing characteristics in such a way that provided a highly personalized vocabulary.

Bartók's recordings, transcriptions, and documentation from Transylvania was meant to catalogue intrinsic music that he hoped might reveal universal musical behavior. His research isolated unique characteristics, but the ultimate compositional purpose was to identify such universal attributes, which, if fully apprehended, might be seen to motivate innate or even atavistic reactions of listener or performer to a composition. Presume, for example, that a theoretical relationship between intrinsic and universal exists, not so much a parallel or reflexive relationship, but rather a derivational one that could reveal general principles, culturally agnostic primitives of musical practice.

The search for such a relationship seems a reasonable motive for carrying on Bartók's mission. Yet five decades after Sachs and ten after Bartók, how likely is it that we'll find isolated regions of indigenous culture in which to explore? The world has changed profoundly in the past century. Whereas before, Bartók could travel to a remote Transylvanian village to study folk music, today, that village very likely has access to radio, TV, and internet, and just as likely, many of its villagers have emigrated. Whereas before mountains and seas may have buffered remote regions, today trains, freeways, and planes have all but eliminated the ability of those geographical barriers to isolate and preserve. An electron, knowing nothing of regional borders, doesn't discriminate which territory it might occupy.

In light of the changes of the past century, while Bartók's ideas remain prescient, new questions arise. For instance, do pockets of unique culture still exist within our modern world? If so, where would one find them and how would one identify them? Are they static or constantly emerging and evolving? Given the growing prevalence of technology and the associated wide-scale dissemination of information, have the cultures of the world become more congruent? Is there a notion of a region as it relates to culture? If, as remote Transylvanian villagers hear Lady Gaga on the radio or YouTube, have we, in aggregate, become less diverse, more dull? Has the advent of the modern era accelerated a Darwinian process of reducing existing diversity and flattening culture? If it has, our search for the universal, or even the intrinsic, may prove more daunting than ever.

1.2 *espressivo* and time

Rhapsody No. 2: lassú és friss

1.2.1 time

Revisiting Bartók's Transylvanian village a century later, we seek which universal attributes that belie seemingly unique cultural characteristics of a particular music. Although we could, like Bartók, start with melodic or rhythmic features, it seems perhaps more fruitful for this research to approach the core issue of *expressivity* by evaluating performance attributes, and in particular tempo and agogics.[2] According to Todd,

> During a performance, a pianist has direct control over only two variables, duration and intensity (Seashore, "The Psychology of Music"). Other factors such as pitch and timbre are determined largely by the composer and the mechanics of the instrument. Thus expressiveness imparted to a performance lies in the departures from metrical rigidity and constant intensity. (Todd, "A Model of Expressive Timing in Tonal Music")

[2] Thiemel defines *agogics* in Grove Music Online:

> A qualification of EXPRESSION and particularly of ACCENTUATION and ACCENT. The qualification is concerned with variations of duration rather than of dynamic level.
>
> A pause of breath of phrasing (suspiratio) is mentioned in a number of organum sources, and in the 16th century the pause (suspirium) was recognized as having affective value. Calvisius recommended delaying or accelerating the beat in connection with the harmony and the sung text (1602). Modifications of the basic tempo seem to have become increasingly common during this period; they are clearly described in Frescobaldi's preface to his first book of toccatas, and are also mentioned by Monteverdi. (Thiemel)

If you accept Todd's point and impose one more constraint, trading the Steinway for a harpsichord with basic disposition, then we are left with *time* as the sole means of expression.[3]

The discussion of musical expression motivates a connate question: what to make of the musical term *espressivo*? The reference to *espressivo* in Grove states "a mark of expression" (Grove). Yet the definition, while concise, has a single reference, namely "see also tempo and expression marks" (Grove). Kolisch observes that "in expressive playing there are rarely any two consecutive bars or even any two consecutive beats at precisely the same tempo" (Kolisch; Fallows).[4] Expressivity, then, is bound to the perceived manipulation of segmented or 'measured' time. It follows that a strict tempo as suggested by a metronome mark becomes controversial.

> In 1826 Beethoven wrote to Schott: 'We can hardly have any tempi ordinari any more, now that we must follow our free inspiration'. The Romantic search for individuality had made the obvious tempo something to be despised. In 1817 he had written to Hofrat von Mosel saying that he wished to discard the 'four principal tempos' (allegro, andante, adagio and presto) and to use a metronome for tempo, but added: 'the words that indicate the character of a piece are another thing... these terms refer actually to its spirit, which is what I am interested in.' (Fallows)

Certainly there is a connection between our sense of 'time' and what Beethoven calls music's 'spirit'. Together they frame our perception of performances of notes and phrases, but also our very understanding of the spirit and essence of a piece. Or more. Fallows notes that Liszt, Bartók, and Kodály used the Hungarian *lassú* (slowly) and *friss* (fast) for music in the folk style, drawing attention, as tempo and expression marks generally do, to particular traditions within which the pieces belong, and to the unique agogic practices associated with those styles (Fallows).

1.2.2 motion

The term *motion* is often used metaphorically to represent the fundamental attributes of melodic shape and of musical performance. The study of musical motion may therefore elucidate intrinsic and universal behaviors. Helmholtz agrees.

[3] Articulation itself, which influences timbre, relates to a performer's control of more refined time-frames. Listen to Bartók's own recording of his "Quasi Pizzicato" from the *Petite Suite*.

[4] Listen, for example, to an early recording of Eibenschutz performing Brahms' Intermezzo Op. 119 No. 2.

> Hence, again, it becomes possible for motion in music to imitate the peculiar characteristics of motive forces in space, that is, to form an image of the various impulses and forces which lie at the root of motion. And on this, as I believe, essentially depends the power of music to picture emotion. (Helmholtz)

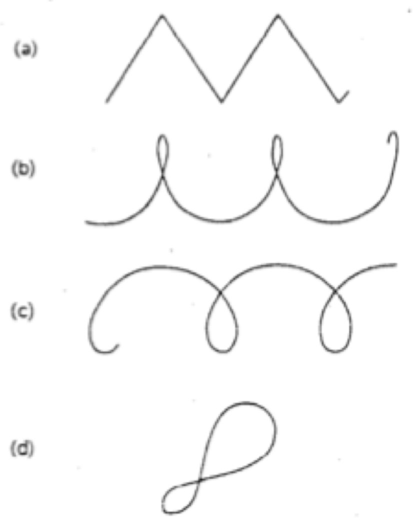

FIGURE 1.1: Truslit universal motion models (Truslit).

In his *Gestaltung und Bewengung in der Musik* published in 1938 and later translated by B. Repp, Truslit argued that motion in music, as conveyed by the performer, actually derives from the inner motion of the performer, as governed by the vestibule, a hypothesis further explored by Todd, Friberg and Sunberg, Repp, Das, Howard, and Smith and others (Truslit; Todd, "A Model of Expressive Timing in Tonal Music"; "The Dynamics of Dynamics: A Model of musical expression"; Repp; Friberg and Sundberg; Das, Howard, and S. L. Smith). Truslit proposes four categories of motion, one 'unnatural' and three 'natural', graphed as in Figure 1.1 and described as follows:

(a) *Straight*, an unnatural motion.

(b) The *open* movement begins calmly, accelerates on the way up, makes a narrow counter-clockwise loop, and decelerates on the way down.

(c) The *closed* movement begins rapidly, decelerates as it reaches the top, then accelerates on the way down, making a larger clockwise loop if it continues into another movement.

(d) The *winding (gewunden)* movement ascends diagonally into a large counter-clockwise loop and descends fairly vertically, making a smaller clockwise loop at the bottom, which leads it back to its origin. (Truslit)

Todd, in his study *The dynamics of dynamics: A model of musical expression*, claims that crescendi and decrescendi detected in performance enable us to identify musical phrases (Todd, "The Dynamics of Dynamics: A Model of musical expression"). Friberg and Sundberg developed a mathematical model that they apply both to final ritardandi in music performance and to runners' decreasing velocity when coming to a halt (Friberg and Sundberg). In their study linking locomotion to music performance they point out that "music is commonly associated with motion. This is manifested by the common use of motion words in descriptions of music, such as lento, andante, corrente (slow, walking running)." Indeed, the relationship between musical time and expression is intimate and profound.

1.3 the elusive search

A kékszakállú herceg vára

The metaphor of musical *motion* thus represents a potential means of describing universal attributes of musical expression, and motion analysis, particularly in terms of a performer's inflection upon a composer's quantized roadmap of a musical score, offers promise in the search for the universal. Nonetheless, while the study of motion in music as a means of identifying universal behaviors has precedent, the search for the isolated Transylvanian village remains challenging, in no small part because we haven't heard from their inhabitants. Todd, in a paper that postulated a link between performed ritardandi and structural boundaries in music, chose an undergraduate student and two concert pianists, respectively, to play Mozart's A Major Sonata K 331, Haydn's Sonata 59 Adagio, and Chopin's Trois Nouvelles Etudes No. 3 (Todd, "A Model of Expressive Timing in Tonal Music"). In the Friberg/Sundberg study on velocity curves during runners' stopping, the sampling of data sources included two male and two female dancers, in both cases professionals, and three musical excerpts of Bach, "determined according to the advice of two professional musicians." (Friberg and Sundberg) Das, Howard, and Smith, in their attempt to validate Truslit's theories on universal, intrinsic motion, observed "five competent performers." (Das, Howard, and S. L. Smith)

Given the objectives of this research, the question is whether these studies aid our understanding of universal behaviors, or merely assess the impact of pedagogy on the style. To what extent did the performers slow down at phrase boundaries because of learned practices? Absent musical indoctrination in performance practices, might there be a tendency for untrained performers to 'rush' through these structural boundaries? If we draw an analogy to the Sachs reference of cultural pollution, one might propose that these studies have simply substituted one source of pollution for another, perhaps a music teacher in place of a Soviet Agent or American oil driller. Even in instances where such studies have applied formal methods and standards, we find examples of bias in the choice of the subjects and note the absence of statistical rigor because of the dearth of data.

Recognizing such limitations, Repp took a different approach in his study of performance practices, collecting a larger sample of performances, analyzing "28 performances by 24 outstanding pianists." (Repp) Moreover, while acknowledging and fully accepting the impact of pedagogy by examining only expert performances, he attempted to identify and segregate the common elements of 'learned practices' of the performances from unique individual expression. He cites Seashore as motivation for his work:

> ...there is a common stock of principles which competent artists tend to observe;... We should not, of course, assume that there is only one way of phrasing a given selection, but, even with such freedom, two artists will reveal many common principles of artistic deviation. Furthermore, insofar as there are consistent differences in their phrasing, these differences may reveal elements of musical individuality. (Seashore, *In Search of Beauty in Music: A scientific approach to musical esthetics* 77)

Building on Seashore's theory, Repp seeks to formalize a dichotomy between 'canonical' and 'individually expressive' performance practices; he proposes that upon identifying the former, the latter, being isolated, become self-evident.

> Though teachers may differ considerably in their methods and goals, and are rarely very explicit about what these are, they are transmitting the unwritten rules (though see Lussy, 1882) of a performance tradition that goes back to 19th century central Europe, where most of the standard repertoire originated. Despite various changes in performance practices during the last 200 years, most of them of a narrowly technical nature, there are generally accepted norms of musical performance, according to which the artist's actions are largely subordinated to the musical structure...
>
> Thus there are two basic aspects of music performance: a normative aspect (i.e., commonality) that represents what is expected from a competent performer and is largely shared by different artists, and an individual aspect (i.e., diversity) that differentiates performers. (Repp 228)

If we accept this distinction, the study of both normative and individual aspects of performances seems valuable, as does research that might distinguish the boundary between the two. Such studies may very well uncover relationships between the performance of music and various factors including language, physical properties, physiological functions, music's underlying structure, or even basic listener understanding of music. But one question regarding such studies remains paramount: whether the normative behaviors indeed follow canonical forms. If they don't, how are we to segregate the normative from the individual (e.g. the "Horowitz" and the "Cortot" factors)?

> Thus, the Horowitz and Cortot factors are not artifacts due to the overrepresentation of these artists in the sample. They represent two true alternatives to the 'standard' pattern of performance timing instantiated by Factor I — alternatives that only Horowitz and Cortot dared to choose in nearly pure form, but that several other pianists incorporated partially into their timing strategies. (Repp 241)
>
> It is probably futile to attempt to characterize the boundaries of acceptable performance practice. They are manifold and are likely to contract and expand as a function of many factors. It is theoretically more parsimonious to conceive of a performance ideal (norm, prototype) that lies at the center of the hypothetical space enclosed by the boundaries. (Repp 256)

How formal, then, is the boundary between pedagogically developed normative forms and individual expression? Chafe, for example, analyzed several recordings of the Charles Ives' *First Piano Sonata* by a single expert, George Barth, a Professor of Performance at Stanford, over a four month period of time (Chafe). Unknown to the pianist, a specific 55-note extract from the fifth movement was compared across five recordings. Figure 1.2 graphs the note onsets and key velocities for all five takes. Per Repp, did Barth's interpretation across the recordings vary from normative to Horowitz, Cortot, or even Kirkpatrick? Putting aside such speculations, the more relevant question for this research, however, is whether such approaches enhance our understanding of intrinsic behavior, or simply inform the study of what Repp calls 'normative' or learned behavior.

For the time being, then, let us exclude experts and 'learned practices' and attempt to find a significant number of "isolated Transylvanian villagers."[5] One source might be children. In order to test a theory on musical training as related to the acquisition of musical skill, Drake and Palmer used both children and adults as subjects in their experiment (Drake and Palmer). Their results (see Figure 1.3) suggest that the least skilled produced the most errors as measured across categories of pitch (both accuracy

[5] The topic of expert performances will return in Chapter 3.

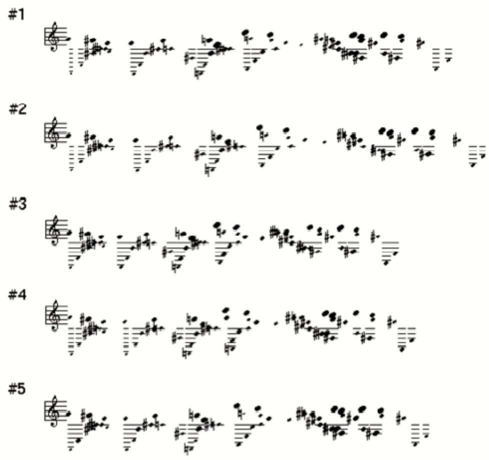

Figure 1: Displayed proportionally, the raw data for note onsets and key velocity shows expressive variations.

FIGURE 1.2: Agogic and dynamic variance in expert recordings of Ives’ First Piano Sonata (Chafe).

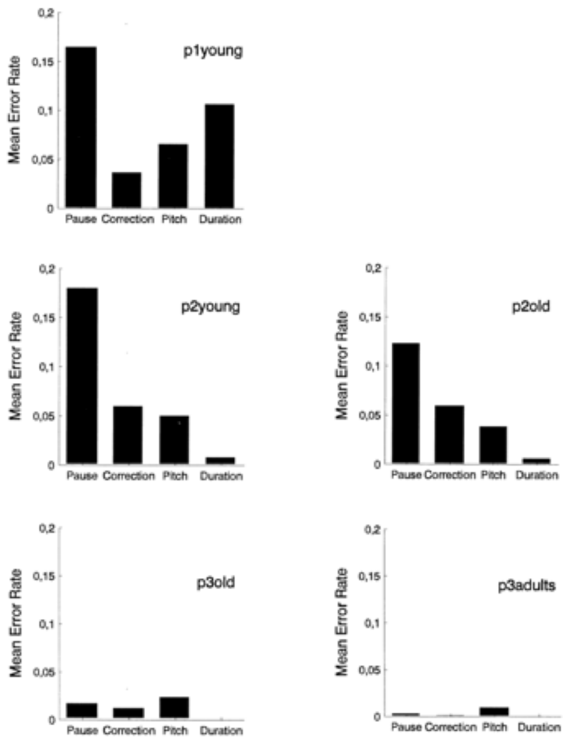

Fig. 4. Mean error rates by error type and group.

FIGURE 1.3: Drake and Palmer study of student performance errata by age/experience (Drake and Palmer).

and error correcting) and timing (including treatment of note durations and pausing). Possibly the most interesting facet of this study is the contrast between p1 (least skilled) and p2 (moderately skilled) groups. For example, the group comprising the least skilled children was the only group where timing errors of note duration were pronounced.

Yet even these children were not entirely unskilled, as Drake and Palmer chose as subjects only pianists with at least three years of training and with sight-reading ability.[6] So the task of finding the "isolated Transylvanian villager" remains, and worse, the requirement of isolation creates a paradox. Per Elwin via Sachs, "one of the most tragic things about contact of the aboriginal with civilization is the destruction of art and culture that so frequently follows." (Elwin and Hivale) Absent such contact with a subject, and absent subjects who demonstrate basic fluency, how exactly would one measure universal musical performance attributes?

Perhaps given this paradox, other researchers have sought to derive universal musical behaviors through the identification of unique cultural attributes. Sadakata, for example, considers the impact of "language" on the interpretation of music:

> There has been little work on culturally specific expressive characteristics in the field of music cognition. One reason is that it is very difficult to measure and to generalize about cultural influence on music performance. Countless factors, such as history, education, social circumstances, and interactions among these factors, contribute to the creation of one's cultural environment. It is certainly a challenge to clarify which cultural aspects influence which features of musical expression. (Sadakata 9)

> There are numerous possible factors that characterize culture, such as language, social structure and history. A factor that is likely to have an influence on perception and production of rhythm is language: Japanese, for instance, is hypothesized to be a mora-timed language in which all 'syllables' have the same duration in speech, while English and French are often characterized as stressed-timed and syllable-timed respectively. Though these distinctions have been criticized (Dauer, 1983; Roach, 1982) and other classifications have been proposed (Wenk and Wioland, 1982), it is clear that these languages differ in their timing structure. (Sadakata 17)

Yet Sadakata chose a set of professionally trained musicians as subjects for her study.

[6] "Child and adult pianists ranging in age (9 - 26 years), training (3 - 15 years) and sight-reading ability learned to perform a novel musical piece in eleven practice trials." (Drake and Palmer)

> Twelve percussionists participated in the experiment. Six of them were professional musicians residing in the Netherlands. The other six were percussion majors (five undergraduates and one graduate student) at Kyoto City University of Arts in Japan. The Dutch and Japanese participants had an average of 21.5 years and 17.8 years of musical training respectively. (Sadakata 19).

Hence, in spite of the clear logic linking musical behavior to culture and language, we are left with a study that makes it difficult to distinguish between culture and training, just as Sadakata indicated in her caveat: "it is very difficult to measure and to generalize about cultural influence on music performance." Indeed, the strictures associated with the use of limited and biased subjects still seem to prevent us from identifying universal musical behaviors.

1.4 inside/out and outside/in

Microkosmos No. 25: imitation and inversion

The search for the remote Transylvanian village continues. Having challenged studies that scientifically analyze intrinsic behaviors as well as those that attempt to link unique cultural attributes such as language to these behaviors, let us now turn to an alternative approach, itself an imitation of Bartók's methods. Bartók's efforts to identify folk music could perhaps be characterized as "inside/out" or "cultural export": identification of the actual source, first, the propagation of that source material into music and its dissemination, second.[7] A variation of this approach more or less inverts Bartók's method and might be characterized as an"outside/in" or "cultural import": begin by transmitting music into a region, then study the propagation and impact of the transmission.

Congruent with Bartók's philosophy on the integrity of the identification of the folk original, this outside/in model might serve to signal intrinsic characteristics within, disparities between, and universals shared by regions, as revealed by their assimilation of "imported" music. If, for example, two distinct regions imported the same song (i.e. they both discovered the song and played it via similar means), and yet if the two regions demonstrated unique interpretations of the song, we might conclude that each region exhibited its own cultural bias. Aarden and Huron affirm this model, combining science with Bartók's study of 'place': "The fact that culture may be correlated with place suggests that one approach to identifying possible cultural features in music is to

[7] Recall the earlier reference, [musicologists were] "reversing the natural order of things, in which practice precedes theory" (Stevens).

look for musical characteristics that change systematically with respect to geography." (Aarden and Huron 170)

The chapters that follow seek to develop this model. Specifically, we create a paradigm in which particular songs are 'imported' into a given region. The inhabitants of that region are then provided simple means of performing these songs in a way that demands no prior training. We record, then analyze the performances of each song, identifying behaviors unique to that region and in contrast to or showing affinity affinity with other regions. Such analysis leverages the work of previous researchers, specifically Sapp who successfully used statistical techniques to identify correlations between recordings as revealed by the use of time in performance (Sapp, "Comparative Analysis of Multiple Musical Performances"). The research will rely on a significant yet non-conventional corpus of music performance data culled from several million performances of piano music on mobile phones (see Chapter 2). As each performance in this corpus is geo-tagged with location data, it will be possible to group performances by region, where the definition of region can be a village, a state or country, or even an arbitrary definition such as people performing above ten-thousand feet in the evening.

1.5 a paradox

Suite, Opus 14

Recall the aforementioned paradox: "One of the most tragic things about contact of the aboriginal with civilization is the destruction of art and culture that so frequently follows." (Elwin and Hivale) By adopting this outside/in model to search for universal musical behaviors, the researcher confronts a problem that results from the complex relationship between technology and the study of music, in particular the study of ethno-music. Bartók was able to find and preserve unique cultural artifacts in remote regions partly because of the absence of technology and its impact on the dissemination of information (and peoples). Had the villages access to the same primitive phonograph technology he used to record their songs in order to preserve, catalog, and publish them, his project may have been impaired. We seek to locate unique regional attributes using technology that, even as it identifies potential answers, will also very likely pollute them.

This paradox has analogs. Here we reference not only missionaries, oilmen, but also ethnomusicologists. And to the point, the research contained in this thesis has possibly traded European colonizers for iPhones. This paradox could be viewed from a different perspective, namely the segregation of 'us' and 'them', the impact of that segregation on 'them', or even our inability to interpret 'them' given our bias, both cultural and professional, toward 'us'. Consider Agawu's discussion of this division:

> One of the most persistent and at the same time controversial dichotomies used by ethnomusicologists is the us/them construct. The construct aims at recognizing the differences between researcher and researched, between 'those who seek knowledge about other musical traditions and those that impart the knowledge', and between colonizer and colonized.

He continues:

> A collective 'us', whether a reference to Westerners, white males, or ethnomusicologists, is no more valid than a collective 'them', which lumps together people with different abilities and levels of knowledge about tradition and culture. (Agawu)

Viewed in this light, any methods that attempt to measure unique regional behaviors could be inherently biased. At some level, by exporting western cultural artifacts into these various regions and then analyzing their performance interpretations, we might be accused of thrusting more of 'us' onto 'them', thereby polluting the very regions we seek to analyze (i.e. the paradox). Or, from a different perspective, because of the dichotomy, it may not be possible to fully gauge and interpret regional reactions to these 'western artifacts'. As such, it's conceivable that our methods yield no insight into intrinsic musical behaviors, but merely describe and potentially measure the viscosity of information flowing into these regions (see Chapter 6). With these concerns in mind, our search for the "remote Transylvanian village" continues.

Chapter 2

The data

Mit látsz? — száz kegyetlen szörnyû fegyver

2.1 culture and land

Eintritt

As Aarden and Huron note, the correlation between culture and place creates opportunities to identify cultural features in music through exploring how musical characteristics vary by geography (Aarden and Huron). The relationship between culture and geographic region is axiomatic. The opportunity exists, therefore, to test hypotheses through applications of this Aarden/Huron model. Yet such research requires the availability of a significant corpus of musical data that includes the geographical source of that data. Moreover, in order to substantiate such hypotheses, a significant amount of data is required; otherwise we are, alas, left with little more than interesting anecdotes.

Helmut Schaffrath developed such a data corpus in the late 1980s (Schaffrath). Identifying and cataloguing over six thousand German folksongs, Schaffrath devised an encoding system to represent the data digitally, the Essen Associative Code for monophonic music, or *EsAC*. *EsAC*'s ASCII representation, while originally written to serve analysis programs running under DOS and designed for a now defunct relational database called *AskSam*, was amenable to reformatting to serve programs both in DOS and UNIX operating systems (Schaffrath and Huron). By 1994 over 14,000 folksongs were encoded in *EsAC* format. In addition to the limitation of monophonic encoding, the data structure assumed quantized pitch and rhythmic data. An intended advantage of *EsAC* was that the format was simple to read and decipher by humans as well as amenable to computer-based analysis. Although not designed to record human performance, *EsAC* does include a field for phrase boundaries. In the decades that followed the creation of the Essen folksong collection, the database grew, and an increasing number of researchers wrote analysis programs that used the *EsAC* format.

Roughly ninety percent of the songs in the initial Essen collection included some type of geographic designation, either a country designation (roughly half of the total collection), a regional level (one quarter of the collection), or a city or village (fifteen percent). The remaining songs (about ten percent) have no regional designation. However the Essen collection has limitations. “There is a certain amount of overlap among categories” (Aarden and Huron 171-172). In other words, several of the geographic designations are not precise. In some instances, regions are referenced by multiple names. In other instances, there are inconsistencies in the format given to place names, etc. These factors are compounded by the relatively small amount of data. These limitation reduce the statistical options available for data analysis, because such analysis requires both a significant amount of data and the ability to detect and account for errata. Of interest, Aarden and Huron propose a more reliable method of identifying the location of each folk song in the Essen collection, namely a GPS tag that associates a given latitude and longitude with each song. Such tags could be precise to a square meter.

Notwithstanding these limitations, researchers have used the *EsAC* to test theories of universal musical behavior. For example, Shanahan and Huron developed a hypothesis that asserted, like Sadakata, a link between language and music, the link in this case being interval compression (Shanahan and Huron; Sadakata). They compared interval compression of German and Chinese folk music, using the Essen data for the German analysis. In this case, they concluded that the hypothesis didn’t hold: the intervals compressed in the German folk songs but expanded in the Chinese folks songs. The existence of a significant corpus of musical data in an electronically accessible format that includes geographical tagging would represent a significant milestone in the analysis of music across cultures.

Inspired by these recent efforts to encode music and tag it with associated geographical indexing, I introduce in this thesis geographically tagged encoding not of rhythmically quantized melodies with generalized rhythmic proportions, but of actual performances replete with temporal nuance, including error and unintentional temporal shaping, as well as for presumed expressivity. My premise is that this provides a unique perspective on ‘universality’ in music.

2.2 the Transylvanian villagers?

Von fremden ländern und menschen

According to Strategy Analytics, a research firm that forecasts trends in telecommunications and other markets, as of October 2012, over one billion people — that is, one in seven people in the world — use a smartphone or related mobile device (Analytics). Strategy Analytics predicts that this number will double by 2015. Research firm IDC confirmed another 200 million devices were sold in the first quarter of 2013 (IDC). Riding this mobile growth — as reported by Fost — as of May, 2013, mobile software developer Smule had over one-hundred million users of their music performance applications (Fost). This constitutes approximately 1.4 percent of the world's population. Collectively these users have created over one billion musical performances. The enormous number of users of these applications, the vast majority of whom have little or no prior training and are broadly distributed throughout the globe, provide a data set of unprecedented richness (Hamilton, J. Smith, and Wang).

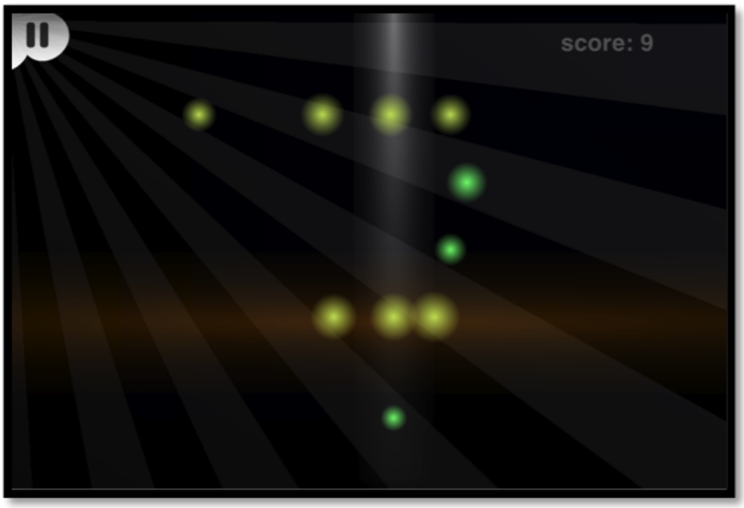

FIGURE 2.1: *Magic Piano* fireflies.

This research examined a subset of the data source generated by the Smule applications, specifically performance data generated by users of the "Magic Piano" mobile application. In this application, users select a song from a catalog and then follow descending fireflies, or cues, that indicate how many 'notes' to tap on a touch-pad screen and when to tap them. Users can make mistakes by tapping in the wrong region or tapping the wrong combination of notes. They control the timing of each note or note combination, dictating the overall tempo for a performance as well as accelerandos and ritardandos (Wang).

The research evaluated performances of over three hundred songs, the vast majority comprised of Western classical and pop songs, but also including some regional folk songs, video game songs, and regional pop songs. A total of approximately eight hundred thousand performances of these songs was analyzed, which comprised roughly six hundred thousand unique performers across over five hundred thousand unique geographic regions. Figure 2.2 shows the location of each unique performance examined in this thesis. The performances occurred primarily during the year 2012. Following Huron's guidance, this data set has been published in open formats and standards to enable other researchers to access it (see Appendix B).

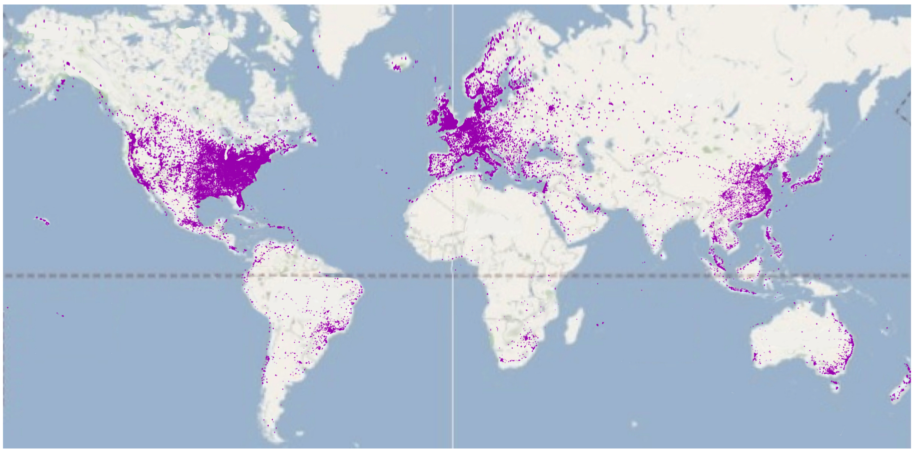

FIGURE 2.2: Location of each performance in the data set.

2.2.1 data source

A *MIDI* file representing the user's interpretation of a given song — their performance — is stored.[1] The captured *MIDI* file also has associated meta-data, including a unique identifier of the user, a stamp of the time and date of the performance, and the GPS coordinates for the mobile device at the time of the performance. Roughly fifty percent of users elect to publicly share their GPS coordinates, and so half of all recorded performances are geotagged.[2] Such geotagging is generally accurate, and in some instances can be precise to a radius of ten meters.

[1] *MIDI* will quantize (i.e. round) notes to the nearest discreet time (in this case millisecond) and loses some fidelity and nuance that might be recorded from an *analog* performance, yet the qualified precision and accessibility of *MIDI* representation would seem to offset these compromises.

[2] The roughly eight-hundred thousand performances in this corpus came from a set of 1.6 million performances, of which half did not contain GPS information and so were pruned from the data set.

2.2.2 potential value of the data source

Aside from the sheer volume of data in this collection, there are several other attractive features:

- the data originates in an accessible electronic format and need not undergo translation or external intervention.
- the data provides relatively precise timings for each note (limited only by the time resolution of MIDI encoding).
- the data is geo-tagged.
- the data is published in open formats including MIDI and JSON.

These features facilitate reliable statistical analysis of attributes associated with performance timings both within and across regions. And the data is "clean" in that it was neither collected, identified, cataloged, translated, nor assembled by a human. The performances themselves occurred on the very electronic device that captured and stored the performance data. Hence the overall accuracy and integrity of the data is extremely high in comparison to data gathered by processes that rely on human intervention.

Since the data is geo-tagged, reverse geo-coding and mapping technology can be applied to identify a place, and "place" itself can be freely defined. This research, for example, examined differences between urban and rural areas, countries, states, and even more arbitrary geographic regions (e.g. the division of earth into every 120 square kilometer region with a known performance). Finally, the accessibility of the application, which uses a novel and somewhat unconventional mechanic, renders the vast majority of performances as amateur in skill. By contrast, most studies in performance to date have focused on expert performance data; hence the fact that this repository of data is amateur could be viewed as an advantage.

On the other hand, although the application affords users an intuitive musical instrument interface, its novelty — the very fact that it is *not* a traditional instrument — may also in some sense be a disadvantage. The mechanic of the application itself is peculiar in that it presents descending 'fireflies' as cues for the striking of notes. The rate at which these 'fireflies' appear, which varies with user behavior, will undoubtably impact a performer's note timings and durations. Hence, comparisons of this data to performances with traditional instruments should be undertaken with appropriate caution. It would seem that most of the insight to be gained from data analysis will involve comparing performances within the data set simply because they all use the same mechanic. Another limitation of the data relates to dynamics, as note-strike velocity is not presently stored in the performance representation.

Throughout this study, aberrant performances were pruned from the analysis based on nonconformance to a standard of deviation (roughly five percent of performances). These included, for example, performances with overall durations at double or half the average for a particular song. At some level, however, the very exercise of removing performances that fell outside the standard of deviation could be construed as controversial, in that it may reduce the level of cultural diversity in the corpus.

Chords can be performed in various ways depending on user preferences and the nature of the song arrangement (that is, whether the notes of chords are aggregated into one, two, three, or four note sequences). While logic in the analysis was developed to 'unpack' chords (i.e. inserting approximate durations for the notes of chords in the sequence of timings) or even prune performances with detected chord preferences, absent such logic, a comparison of two performances might yield different results depending on whether both users shared the same preferences for chord presentation.[3] Indeed, the technique and preference chosen to perform chords may in itself indicate a regional orientation and preference. Moreover, if a given song always aggregates chords in the arrangement, two performances of the same arrangement that demonstrate different interpretations are of course noteworthy, etc.

We should consider other questions regarding the corpus. What set of users has elected to purchase a smartphone or tablet, and to what extent might these demographics skew results, particularly in remote and lesser developed regions? What is the impact of user preference in purchasing and performing songs, given that most songs require either purchase with currency, or access earned by progressing through the product's levels, or by watching advertisements? Finally, although observed usage suggests that the instrument interface is remarkably intuitive and its use demands little, if any, documentation, it should be noted that, until recently, the only text included in the user-interface was in English.

2.2.3 the preponderance of data is derived from amateur performers

As mentioned above, perhaps the most novel aspect of the data set amassed for this research is that it comprises performances that are predominantly done by unskilled

[3] In the case of chords, the analysis often ran both comparisons: one with chord preferences pruned and/or unpacked, and the second with no pruning or unpacking. To date, the removal of such preferences has not generated a significant difference in the comparative analysis, perhaps reflecting the limited use of such preferences, and/or a congruent but limited cultural bias towards such preferences.

performers. The class of devices used to capture these performances — multi-touch smart phones and tablets — did not exist until recently. The *Magic Piano* application itself first became available to consumers in March of 2010, coinciding with the introduction of the iPad by Apple Computer. Moreover, the vast majority of these performers are amateurs in the literal sense of the word, playing music for their own pleasure. Some might even describe their activity as *game playing* rather than as *music making*. If universal tendencies of temporality in music exist, they should reveal themselves here without the biases of any formal music education.

2.3 a model for motion

Andante con moto — Allegro di molto

Different models are possible to measure and analyze motion within a musical performance depending on what information has been captured. Das, Howard, and Smith, in their derivative study of Trustlit, calculated the variance of note duration over time, or $dx(t)/dt$, where dx reflects the difference in sequential *MIDI* note-onset events (Das, Howard, and S. L. Smith; Truslit). The limitation of this model stems from its inability to actually measure changes in time over increments of time. Friberg/Sundberg and Repp instead measure time changes in tempo relative to note in a score, or dt/dn, where, like Das, dt, reflects the difference in sequential *MIDI* note-onset events, and dn reflects the sequential progression through a set of notes in a piece (Friberg and Sundberg; Repp). The note index simply indicates the notes of a piece in sequential order, as illustrated below for *Twinkle, Twinkle Little Star* (see Figure 2.3).

FIGURE 2.3: Note score position (index) values for “Twinkle”.

This research adopts the convention of evaluating time relative to score note positions (or a note index), but proposes an alternative model to better convey tempo changes and agogical events of a given performance or, more likely, a set of performances. In this new model, the tempo at a given note index is calculated relative to a baseline, which could be a representative performance, a fixed tempo, etc. This relative definition of tempi offers a more practical measurement of time, as concepts of fast or slow only make sense from a given, or conventional perspective.

First, note durations are calculated by computing the difference in sequential note onsets (e.g. in the case of a *MIDI* format, note-onset events). Hence duration, D, is defined as

$$D = N_i - N_{i-1} \tag{2.1}$$

where N_i represents the time of the onset of a given note by score position index i. To indicate tempo, the note duration, D, is summed through note index, i

$$T_i = \sum_0^i D_i = \sum_0^i N_i - N_{i-1} \tag{2.2}$$

T_n, then, would give the tempo at final note index n, where note indexes, i, for a given piece range from $0 \leq i \leq n$.

Having calculated the tempo at a given note position, T_i, and the overall tempo, T_n, the variance of tempo over time, as indicated by note indexes, can then be generated by calculating the difference of $T(j)_i$ for performance j to $T(B)_i$ for a baseline performance, for example a fixed tempo. If we substitute for $T(j)_i$ the average tempo for a set of j performances at note index i, or $T(Ave)_i$, and subtract this aggregate average from the baseline tempo, or $T(B)_i$, we get the *a*verage *a*ggregate *t*empo or AAT at note index i.

$$T(Ave)_i = \frac{\sum_0^j T(j)_i}{j} \tag{2.3}$$

$$AAT_i = T(B)_i - T(Ave)_i \tag{2.4}$$

Hence if the AAT_i were negative for note index i, the average tempo for the set of performances would be slower than the baseline, as the average note duration would be larger (i.e. they are of longer duration and hence slower). If the AAT_i were positive at a given note index, the average tempo for a set performances would be faster than the baseline. Of course by this definition, the set of performances could be of size 1 to evaluate the tempo of an individual performance.

Graphing aggregate time, AAT_i, over a sequential series of note indexes then gives you the *slope* of a tempo of a performance, or $d(AAT)/dn$, which henceforth will be referenced as the *slope* of the AAT. The slope at a given note index i is calculated as follows:

$$Slope_i = AAT_i - AAT_{(i-1)} \tag{2.5}$$

This slope indicates acceleration, deceleration, and associated inflection points. If this slope is positive, the AAT is accelerating at this index relative to a baseline; if negative, it is decelerating, while a *slope* of 0 would indicate the AAT has not changed relative to the baseline — they are the same tempo at this note index. Inflection points in the slope therefore indicate changes in acceleration. Like note durations, all units including the AAT and slope are given in milliseconds.

It will also prove useful to evaluate a tempo relative to the average tempo of all performances of a song. To do so, tempo for a given song, T_n, is simply divided by the AAT_n — the average aggregate tempo at the final note index (n) of the song.

$$Tempo = \frac{T_n}{AAT_n} \tag{2.6}$$

The resulting $Tempo$ is given as a percent relative to the average tempo of all performances for a given song, where values above 100% indicate slower tempos and values below 100% indicate faster tempos, namely those performances whose sum of note durations is shorter than average. The potential advantage of this definition is that the same calculation can be used for multiple pieces, thereby enabling aggregate comparisons across pieces and genres of music.

2.4 example application of model

Im legenden-ton

The model for motion described above can also be applied to aggregate data. In this case, data from approximately 23,000 world-wide performances of “Twinkle” was collected and graphed. Such data represents a subset of the performances analyzed and published for this thesis, and could be viewed as representative of the broader corpus. Figure 2.4 graphs the actual location of each performance. Geographically, these “Twinkle” performances were distributed as indicated in Table 2.1. When compared to the 760 largest cities in the world (see Figure 2.5), approximately 57% percent of all “Twinkle” performances took place within an urban region, with over half occurring within Asia.

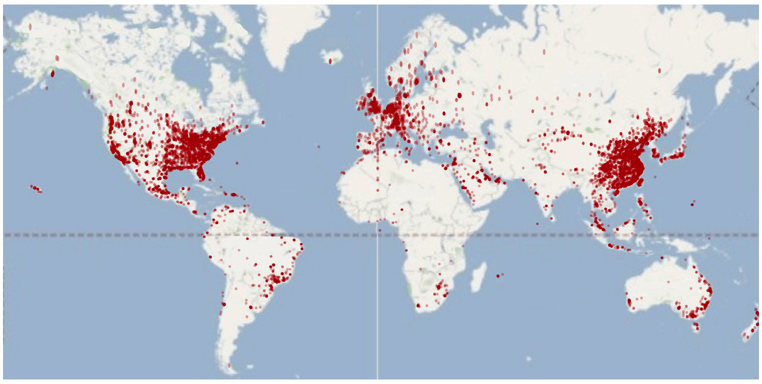

FIGURE 2.4: Location of 23,000 "Twinkle" performances.

Region	Count	Urban Percentage
Asia	12282	69.5
Europe	1440	40.5
N. America	6998	34.8
S. America	1152	72.7
Africa	232	71.1
Australia	853	61.5

TABLE 2.1: Geographical distribution of performances of "Twinkle".

FIGURE 2.5: Top 760 urban centers.

Applying the model for motion to this data set, the AAT for the performances relative to a baseline performance was computed and then graphed, producing Figure 2.6. The slope of the graph indicates varying degrees of acceleration at different moments in the song. At periods of deceleration, inflection points designate a change of slope.

The X-axis of this graph shows the note index, or dn, while the Y-axis indicates the average note duration in aggregate, or da, in milliseconds. The graph of the resulting data then provides a curve that indicates *aggregate average tempo (AAT)* at a given note index.

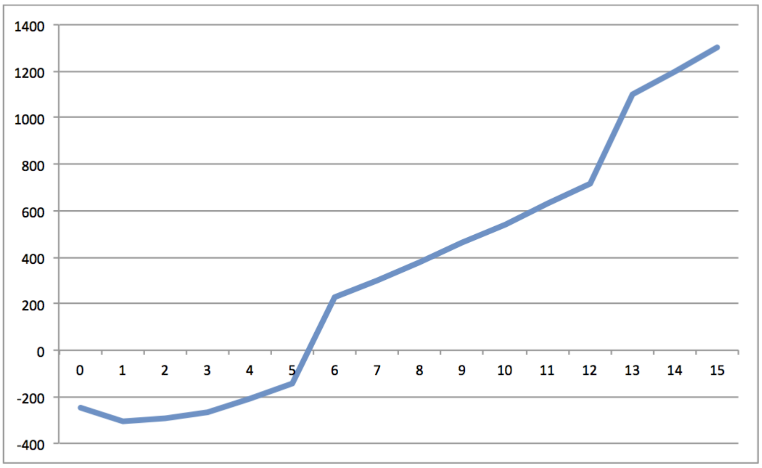

FIGURE 2.6: Slope of 23,000 performances of "Twinkle". The x-axis gives the note index, while the y-axis shows the time in milliseconds of the *AAT* at a given note index. Subsequent graphs of slope will follow the same conventions for the axes.

To interpret the above graph, the 23,000 performances on average started slower than a baseline tempo, which in this case was a fixed tempo. At note index 2, we notice an inflection point where the *AAT* accelerates but moderately. At note index 6, the first half-note in the piece and a phrase[4] boundary (where the melody arrives on the fifth degree of the scale, or V), the *AAT* has largely "caught up" to the baseline. Yet the *AAT* accelerates significantly through the phrase-boundary: the time for the half-note has not been honored. Upon reaching note index 7, the beginning of the next phrase (a melodic consequent), the AAT is almost 400 milliseconds "too fast". A similar inflection point with acceleration exists at the next phrase boundary, or note index 13.

The above example illustrates the application of the proposed model for motion to a sample of the mobile performance data, yielding some general insights into aggregate or even universal musical behaviors in the context of this song. For example, in contrast to previous theories of skilled performances (e.g. Repp, Todd, Friberg and Sundberg) yet somewhat consistent with those of unskilled performances (e.g. Drake and Palmer), this analysis suggests that amateur performers don't pause at phrase boundaries; in fact it suggests they actually accelerate (Repp; Todd, "A Model of Expressive Timing in Tonal Music"; Friberg and Sundberg; Drake and Palmer). Chapter 3 delves deeper into the model with a view towards universal behaviors. However it is not until Chapter 4, where (as Huron proposes) regions are compared, that the actual *outside/in* method is applied to the data, thereby affording more opportunities to identify intrinsic musical behaviors (if they exist).

[4] Here we use the term phrase to describe the consecutive melodic notes that, if sung, could be taken in one breath, which together form a musical thought. Arguably we could use the term subphrase in place of phrase in this example.

Chapter 3

Universal behavior

Ó be sok kincs

In some century to come, when the school children will whistle popular tunes in quarter-tones — when the diatonic scale will be as obsolete as the pentatonic is now — perhaps then these borderland experiences may be both easily expressed and readily recognized. But maybe music was not intended to satisfy the curious definiteness of man. Maybe it is better to hope that music may always be a transcendental language in the most extravagant sense. Possibly the power of literally distinguishing these "shades of abstraction" — these attributes paralleled by "artistic intuition" (call them what you will) — is ever to be denied man for the same reason that the beginning and end of a circle are to be denied. . . .

Beauty in music is too often confused with something that lets the ears lie back in an easy chair. Many sounds that we are used to do not bother us, and for that reason we are inclined to call them beautiful. Frequently, . . . when a new or unfamiliar work is accepted as beautiful on its first hearing, its fundamental quality is one that tends to put the mind to sleep. . . .

> We see him standing on a summit, at the door of the infinite...
>
> Charles Ives

3.1 $E = mc^2$

Thoreau

> To compare two performances, three performances are required: the reference and the two inputs. (Chafe)

The study of tempi and agogics of geo-tagged music performances on mobile devices has the potential to inform our understanding of universal musical behaviors. Given the expressive potential of time, particular questions pertain to how amateur performers might treat time. Do they accelerate and/or decelerate? If so, do the moments of change pertain to the structural segmentation and boundaries of a piece, or do they seem random? Perhaps cases of radical changes in tempi correspond to changes in texture, rhythm, harmony, or even physiological challenges like a leap in melody or a complex chord. Is it possible that the way amateurs treat time correlates with 'place'?

One challenge for all such questions is the relativity of perception and its dependence on context. As we noted above, fast and slow make sense only in relationship to what is perceived as normative, which may be conventional. For if we compare two performances and deem one fast, might it be that the other one was slow? Recognizing the importance of relativity in the perception of time, the use of a third performance or set of performances as a baseline allows a more coherent approach to comparative analysis. Yet questions persist regarding the baseline itself. If, for example, we evaluate a substantial number of performances relative to a baseline tempo, should this tempo be fixed, or should it follow conventions of ritardandi and accelerandi; and if the latter, whose conventions? Alternatively the average duration in aggregate of all performances could be a baseline, or as another option still, all performances could be compared to a single performance.

The other side of this challenge of relativity, of course, is the opportunity to use relativity itself to establish a baseline through relational invariants. If the comparison of several performances reveals a group of shared behaviors, it's possible normative behavior has been discovered within this group. This group can then serve as a baseline against which others might be analyzed. Correlative analysis can also be applied to such groups, both normative and otherwise, ascertaining potential relationships to secondary variables including 'place', which in turn might advocate a source of intrinsic behaviors. Herein lies the basic premise supporting this thesis. This chapter will look for normative behaviors identified through comparisons to various baselines. Chapter 4 will go beyond normative and look to correlate behavior with 'place'.

3.2 Truslit and amateurs

Serenity Now

We will first use Truslit's models for universal motion in music as baselines for analysis, and will seek to identify whether these models represent normative behavior, which we could infer as universal behavior. Do amateur performers' treatment of time conform to his model for motion in music (Truslit)? Other researchers, including Das, Howard, and Smith, have tested this model, albeit with only a handful of skilled performers (Das, Howard, and S. L. Smith). But what about amateur performances en masse?

FIGURE 3.1: Note score position (index) values for "Twinkle".

Recall that Truslit proposes that motion in music, as conveyed by the performer, actually derives from the inner motion of the performer as governed by the vestibule. "The novelist Jean Paul came close to the truth when he said: 'Music is an invisible dance, just as dance is inaudible music."' (Truslit 267) He proposes that this inner motion may follow one of three natural arcs, namely 'open', 'closed', and 'winding' (see Figure 1.1), "the open scale being the fastest, the winding the slowest." (269) He then provides examples of how such motion might characterize ascending and descending scales and arpeggios, and also different sections of known music including excerpts from Wagner's *Tristan and Isolde*. If one were to crudely summarize his model, and it must be somewhat crude as the 'model' itself is abstract if not nebulous, it would be to recommend that music's motion should be natural and continuous, with undulating transitions between acceleration and deceleration.

As Truslit's 'arcs' initially described the motion of ascending and descending patterns, we turned again to the song "Twinkle Twinkle Little Star" because of its clear patterns

FIGURE 3.2: Truslit 'open' motion.

of ascent and descent.[1] To simplify the model, only excerpts from the first 42 notes[2] of the song were analyzed (one full verse, see Figure 3.1), because subsequent to that point, the texture migrates from simple melody to more complex chords. Two of the three Truslit models were tested: 'open' (see Figure 3.2) and 'closed' (see Figure 3.3), using the 23,000 performances referenced in Chapter 2. To perform the tests, we created hypothetical AAT values (see Equation 2.4) for each of the Truslit models for the first two phrases of the song. The slopes (see Equation 2.5) of these models were then graphed (see Figure 3.4). The baseline used for the comparisons was a fixed tempo calibrated to match the average tempo of the performances, or roughly a quarter-note at 160 bpm.[3]

FIGURE 3.3: Truslit 'closed' motion.

As shown in the graph, the 'open' model initially decelerates, and then at note index 4 accelerates, which increase coincides with the descent of the melodic line (note indexes 7 - 8). At note index 10, the 'open' AAT decelerates, corresponding to the melodic arc returning to its origin, the tonic. The 'closed' model begins with immediate acceleration that decreases as the melodic crest is approached and reached (note indexes 5 - 7), at which point it decelerates. The periods of most rapid acceleration and deceleration in the 'closed' model occur at the initial and ending points of the two phrases (note indexes 1 and 13 respectively). In contrast, the periods of most rapid acceleration in the 'open' model occur at the melodic crest and descent (note indexes 4 and 7 respectively), while the deceleration is deferred (note index 11).

[1] The song "Twinkle, Twinkle Little Star", as commonly referenced throughout this study, is likely a French folk song in origin, "Ah! vous dirai-je, Maman". The English lyrics (and title) were subsequently adopted in several English-speaking regions primarily in the context of children's music (Wikipedia, "Twelve Variations on "Ah vous dirai-je, Maman"").

[2] The arrangement of "Twinkle" has the first verse in single notes (the first 42 notes) and then the second verse in a progressively advanced chordal arrangement.

[3] As a point of reference, the AAT for all "Twinkle" performances in Figure 3.4 was recalibrated to the baseline tempo (with the quarter-note at 160 bpm) and appear as 'Average AAT' in the graph.

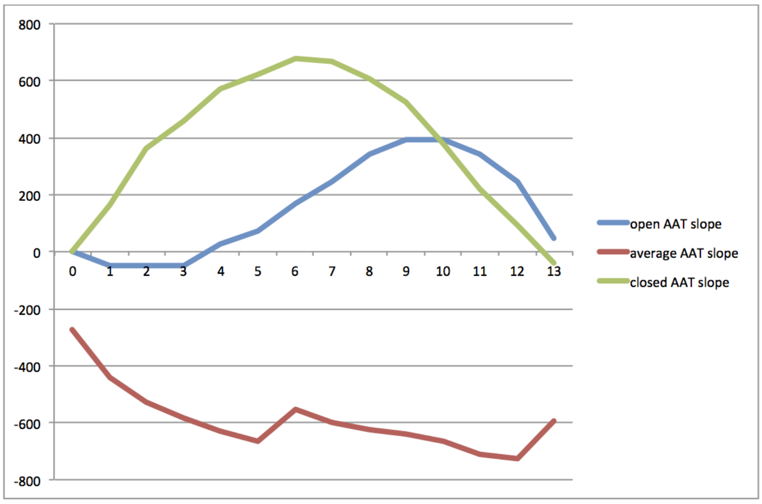

FIGURE 3.4: An AAT representation of two Truslit models for "Twinkle".

Having created an approximate representation of the Truslit 'open' and 'closed' models, we then compared the note durations of each of the referenced performances to these models using the *Pearson correlation coefficient* (see Appendix A) in order to ascertain to what extent amateur performances conform to either an 'open' or 'closed' contour. As the *Pearson correlation coefficient*, which ranges from $1.0 \leq 0.0 \leq -1.0$, will advance towards 1.0 when the comparison between two sets of numbers approaches identity (and conversely -1.0 when the comparison yields a negative correlation, with 0.0 suggesting no correlation), all performances with a correlation of $r \geq 0.7$ were deemed similar to the 'open' or 'closed' models.

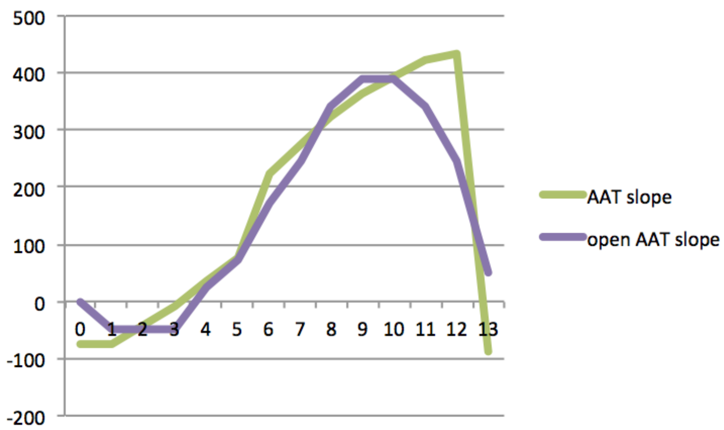

FIGURE 3.5: Truslit 'open' model compared to similar performances' AAT.

Of the 23,000 performances examined, 7.8% or 1,797 performances fit the the Truslit 'open' model while 4.5% or 1,024 performances were similar to the 'closed' model. Figure 3.5 shows the slope of the AAT for the 1,797 'open' performances graphed alongside

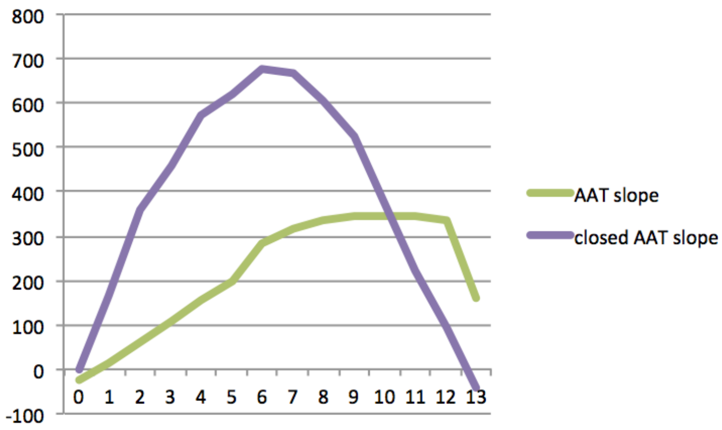

FIGURE 3.6: Truslit 'closed' model compared to similar performances' *AAT*.

the Truslit 'open' model. Figure 3.6 shows the slope of the $1,024$ performances deemed similar to the 'closed' model. Note that Truslit predicts that an 'open' motion would align more closely with the melodic arc of the opening of "Twinkle". In this case, a significantly larger number of performances in fact do. Moreover, the slope of the *AAT* of these performances fairly closely tracks the slope of the model. By contrast, the slope of the 'closed' performances doesn't. Anecdotally, representative 'open' and 'closed' performances are referenced; the 'open' performance took place in Cumberland, Indiana, USA while the 'closed' performance occurred in Aliağa, Turkey. Their respective *AAT* slopes are graphed in Figure 3.7.

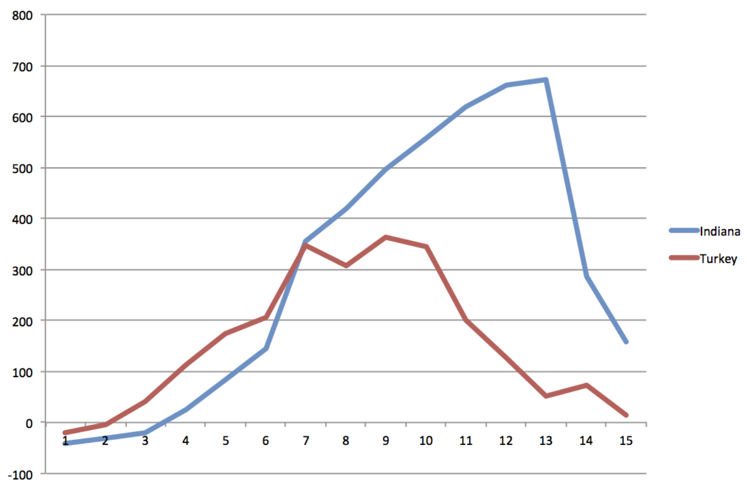

FIGURE 3.7: Slopes of an 'open' and 'closed' performance from Indiana and Turkey.

Perhaps of more significance, given that less than 10 percent of the performances were deemed similar to either model (subject of course to the methods and data described), such Truslit models of motion do not appear to manifest *naturally*: the small percentages contraindicate normative behavior. In fact, if the analysis extends to include the performance of both versus of "Twinkle" (noting, again, that the second verse repeats the melody but with an increasingly complex chord arrangement), the set of performances within each group did not exhibit strong likenesses to each other. To gauge the degree of similarity between performances within each group, we used the average *Pearson correlation coefficient* or *A-Pcc*.

3.2.1 *A-Pcc*

We calculated a group's *A-Pcc* by evaluating the *Pearson correlation coefficient* of every performance to every other performance in the group, and then dividing by the number of total comparisons (see Appendix A.2). If there were a high degree of congruence across performances within a group, the *A-Pcc* would approach 1.0. In the case of non-congruence, the *A-Pcc* would approach 0.0, and −1.0 for negative congruence. Hence the *A-Pcc* becomes a useful metric to compare the degree of interpretative similarities between different groups of performances.

Surprisingly in this case, the overall set of performances of the entire song had a higher *A-Pcc* than either the groups of 'open' or 'closed' performances. Specifically, the *A-Pcc* for all performances (23k) was 0.205; the 'open' group (1.8k) *A-Pcc* was 0.170 versus 0.174 for the 'closed' group (1k). In other words, while their initial melodic arcs conformed to Truslit 'open' and 'closed' models, as the song progressed into a less simplistic form, their conformance to either model degraded substantially. In essence, the modeled behaviors seem neither intrinsic nor universal. In his translator's epilogue of the Truslit paper, Repp proposes Truslit was performance oriented. "Most likely, he was a music teacher rather than an academic; his theories seem to derive from an intensive active involvement with music." (Truslit) Related passages from Truslit's work emphasize the significance of this perspective.

> The unnatural temporal precision of the "straight" scale may be noted. Such mechanical exactness suppresses the inner life and restricts the musical content to the sonic appearance; it has no expressive effect on the listener. In the various forms of motion, on the other hand, if the agogic subtleties occur in the right manner, the tones also seem to arrive with great temporal precision, but with a liveliness that contrasts starkly with metronomic playing. (Truslit 269)

And later

> In some musical examples, it is fairly obvious what this original motion must have been (e.g. the open movement of waving in the Tristan example). In most cases, however, the determination is not so simple and requires a very fine sense for motion. (Truslit 271)

In other words, the natural motion Truslit describes is *learned*. It's not innate. At some level, therefore, the study of Truslit's models, per Repp, may help identify canonical forms of pedagogy, but offers less promise as applied to the study of intrinsic behaviors of amateur performers.

3.3 phrase boundaries

Putnam's Camp, Redding, Connecticut

Drawing once again upon the presumed relationship between language and music (see Section 1.3), intuition suggests that performers, like speakers, delineate phrases through the use of agogics, allocating more time to indicate the boundaries of phrases. "One of the most well-documented relationships is the marking of group boundaries, especially phrases, with decreases in tempo and dynamics (Henderson 1936)." (Palmer) Several researchers of motion in musical performance have looked to link agogics to structure, including Todd, Friberg and Sunberg among others (Todd, "A Model of Expressive Timing in Tonal Music"; "The Dynamics of Dynamics: A Model of musical expression"; Friberg and Sundberg). "This effect [phrase final lengthening] is simply the tendency to slow at the end of a single motor action or sequence. It has been known and well documented for some time (Todd, "A Model of Expressive Timing in Tonal Music" 34).

Todd developed a mathematical model that linked such ritardandi to the structural elements of a piece. His model analyzed two experts' performances. "This hypothesis agrees, in principle, with Cone (1967), when he says 'a valid performance depends primarily on the perception and communication of the rhythmic life of a composition. That is, we must first discover the shape of the piece. . . and then try to make it as clear as possible to our listeners'." (Todd, "A Model of Expressive Timing in Tonal Music" 40) Friberg and Sundberg took it a step further in their study (1998) by constructing a physical model that drew correlations between such ritardandi in music and the motion of a dancer stopping, but as mentioned above, yet again with a handful of expert pianists and dancers.

In her survey on musical performance research, Palmer confirmed performers' use of ritardandi around phrase boundaries, but qualified the association to apply primarily to trained musicians:

> Expressive timing patterns decreased when pianists attempted to play without interpretation, and these patterns increased in exaggerated interpretations, similar to other findings of modulations in expressive level (Kendall & Carterette 1990, Seashore 1938). Further studies indicated that the expressive timing patterns increased from novices to experts, increased during practice of an unfamiliar piece, and changed across different interpretations of the same piece performed by the same pianist (Palmer 1988)...
>
> Musical experience enhances both performers' use of expression to emphasize interpretations and listeners' ability to identify interpretations and expressive aspects of performance (Geringer & Madsen 1987, Johnson 1996, Palmer 1988, Sloboda 1985a). (Palmer)

Yet in contrast to some of her peers, Palmer, citing Dunsby, claims that subtlety exists in the link between musical structure such as phrase boundaries and a performer's use of agogics:

> The field of music analysis offers various explanations for the content of a given composition. For instance, a piece can be viewed as a hierarchy of part/whole relationships, as a linear course that follows the harmonic tension, or as a series of moods that result in a unity of character (Sundin 1984). However, music analysis does not indicate how a performer actually produces a desired interpretation (Dunsby 1989). (Palmer)

In fact, Palmer, citing Clarke, proposes that expert performers may imitate agogics, obfuscating musical structure:

> Evidence that performers can imitate expressive timing patterns that have an arbitrary relationship to the musical structure suggests that performance expression is not generated solely from structural relationships (Clarke 1993). (Palmer)

Truslit goes further and suggests that few performers possess the skill to shape music:

> It is commonly believed that performance relies on "inspiration", "intuition", "feeling", or the like. However, these subjective states must not guide the recreating artist, because they do not give access to the original motion. Rather, the original motion may elicit them. Unfortunately, there are few artists today who can shape under the influence of inner motion. (Truslit 274)

3.4 expertise required?

The Unanswered Question

The ability to use agogics to demarcate musical structure may very well remain in the province of the expert. Even so, while the preponderance of research finds that learned behavior is required, the delineation between normative and individual expression lacks clarity (see Chapter 1.3 and the discussion of Cortot, Horowitz, and Barth).

Notwithstanding such inherent nuance among expert performances, what about amateurs? By analyzing amateur performances, is it possible to confirm the theory of agogical enunciation of musical structure as an expert-only trait? Or do amateur performers intrinsically identify musical structure, for example through the use of ritardandi?

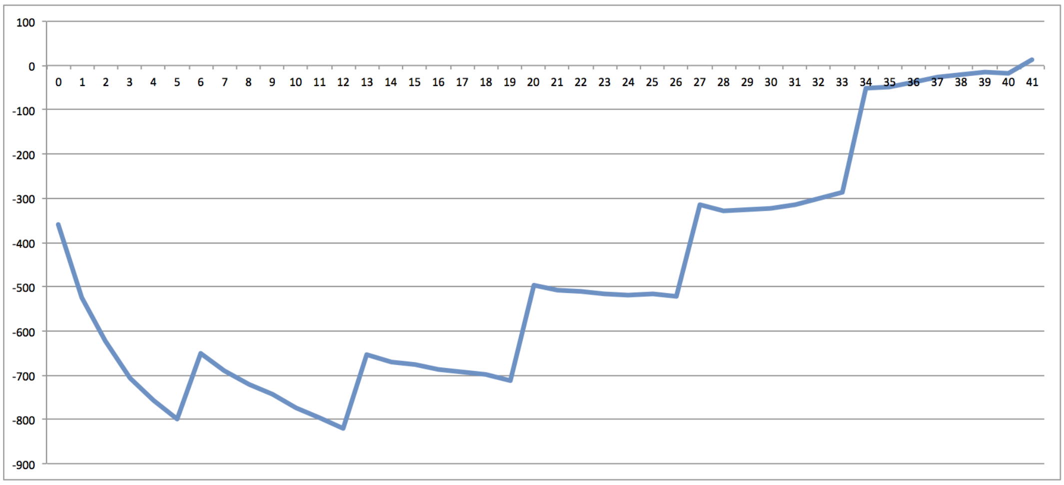

FIGURE 3.8: Slope of 23,000 performances of "Twinkle".

More simply, do performers actually slow down at phrase boundaries? If so, do they decelerate because they learned this behavior or because such behavior is innate? In the case of the amateur performers evaluated for this thesis, generally they do not. Figure 3.8 shows the slope of the *AAT* through note index 41 for "Twinkle" once more. Comparing the slope to the score (see Figure 3.1), several successive inflection points exist, namely note indexes 6, 13, 20, 27, and 34. Each inflection point represents a moment of acceleration in the graph; each also corresponds to a phrase boundary. As demonstrated by the graph, rather than slow down at phrase boundaries, on average these performers sped up.

While the proclivity to accelerate seems strongest at the phrase boundary, a further scrutiny of performances of other songs reveals a related correlation of acceleration to note duration and not musical structure. In "Twinkle", all phrases end with a half-note versus a quarter note, and so each inflection point is also homologous with these

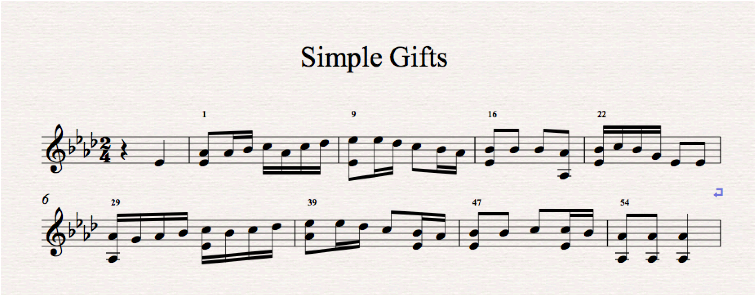

FIGURE 3.9: Note indexes of "Simple Gifts".

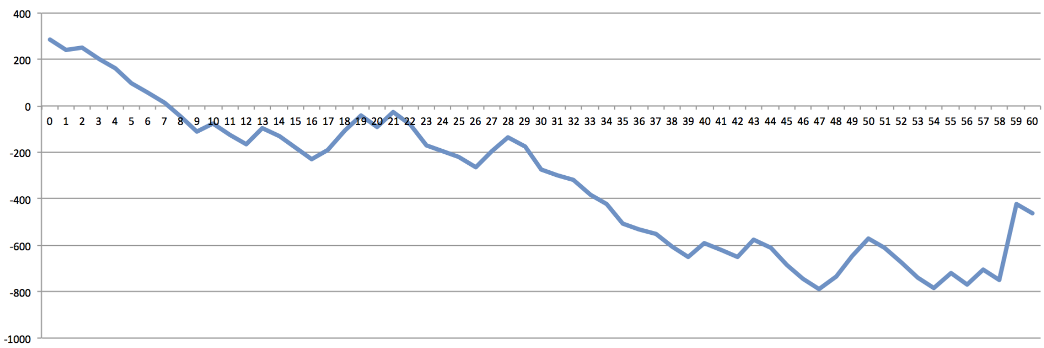

FIGURE 3.10: Slope of 4,047 performances of "Simple Gifts".

half-notes. Performers, on average, are 'cheating' the duration of virtually every note of longer duration. The particular mechanic in the *Magic Piano* product creates the perception that notes following notes of longer duration could be played sooner. While the performer determines when to strike the next note(s) and ultimately controls the timing, the impact of the mechanic cannot be dismissed, as perceptions it encourages may materially affect timing. This fact may prejudice any conclusions about intrinsic behaviors, and particularly if our data were compared with data from performances that don't share this mechanic. Yet when compared with like performances (e.g. two performances that use this same mechanic), our data provides potentially instructive insights into innate tendencies, especially when such performances are then correlated to a secondary variable such as 'place' or even a repeat performance. Considering the potential influence of the product mechanic, and notwithstanding the opportunity to use relational invariance to eliminate or even exploit the impact of this factor in this study, the analysis of other songs including those where phrase boundaries aren't associated with longer note durations will be constructive.

Figure 3.10 graphs the slope for the *AAT* of 4,047 performances of "Simple Gifts", another Western folk melody. The slope, like that of "Twinkle", confirms that performers typically accelerate through each note of longer duration. In "Simple Gifts", however, the rate of acceleration is offset somewhat by extra time taken to manage the complexity of chords, in this case dyads (see Figure 3.9). This latter correlation between time and complexity for amateurs certainly seems obvious; it also has precedent in research:

> Analyses of piano performances indicated that chord errors occurred more often in homophonic styles and that single-note errors occurred more often in polyphonic styles, which suggests that the relevant musical units change across different musical contexts (Palmer & van de Sande 1993, 1995). (Palmer)

The amateur, it would appear at least in terms of an initial performance of a song, has limited conception or perhaps capacity to develop structure through agogics; instead, their focus applies to the basic process of striking the correct sequence and combination of notes.

FIGURE 3.11: Note indexes of "O Holy Night".

In another example, the Western Christmas song "O Holy Night", Figure 3.11 shows the score with associated note indexes while Figure 3.12 gives the slope of performances. The same behavior exhibited in "Simple Gifts" applies to the instances of deceleration in this piece: they are linked to more complex moments, namely cases of dyads versus single notes. At the most significant phrase boundary in the strophic form where the first verse repeats, or bars 10 and 11 corresponding to note indexes 75 - 87, the period where we would expect an expert to decelerate to denote the structure, in fact the unskilled performers accelerate. In this case, the acceleration seems linked to the relative

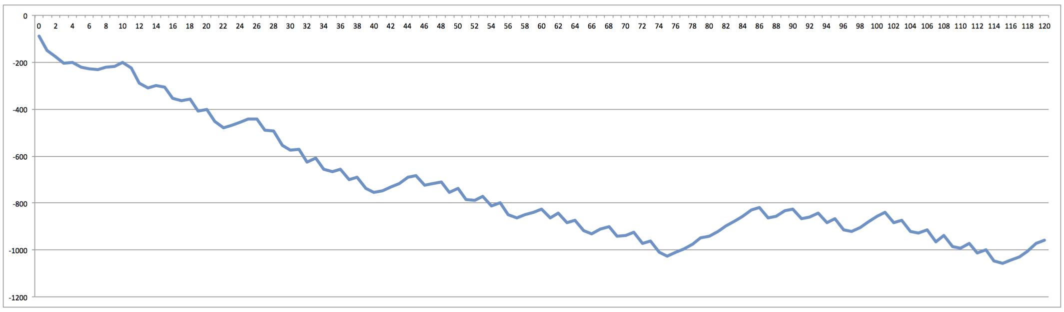

FIGURE 3.12: Slope of 4,741 performances of "O Holy Night".

simplicity of the structure at these bars (single notes versus chords), and in variance to the first examples ("Twinkle" and "Simple Gifts") where accelerandi instead indicated associations to longer note durations. That is, absent notes of longer duration at the boundary of the phrase (and therefore absent the inducement to accelerate via the mechanic of the product), the unskilled still accelerates through a significant phrase boundary, here the end of the first verse.

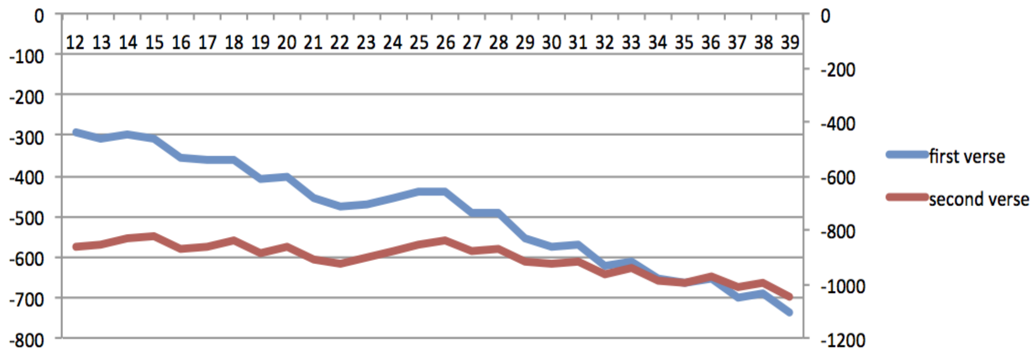

FIGURE 3.13: Slope of bars 3 - 4 and 12 - 13 of "O Holy Night"
(note indexes indicated for bar 3).

The strophic form of "O Holy Night" presents an opportunity to examine how amateur performers treat time in the second instance of the same verse within the same performance. While the verses vary,[4] the first four bars of each verse (bars 3 - 6 and 12 - 15 respectively) are identical. Figure 3.13 graphs the slopes of the AAT of both bars 3 and 4 and bars 12 and 13, where the $y-axis$ has been calibrated to the scale of each verse, namely bars 3 and 4 for the left axis and bars 12 and 13 for the right. As the performers in aggregate decelerate through the first verse, by the time they commence the second verse, their AAT has slowed by approximately 800 *milliseconds*, almost a

[4] The first verse moves harmonically through the subdominant to the tonic, then dominant before arriving on the tonic. The second verse replicates this harmonic motion through the subdominant, but then departs from the first verse (at bar 16) and goes through a secondary dominant to arrive at the mediant.

	Performances	Three or more repeats	Percentage
"Twinkle"	26867	510	1.90%
"Amazing Grace"	22451	481	2.14%

TABLE 3.1: Percentage of performers who repeated three or more times.

full second slower than the baseline tempo. Hence the second $y - axis$ was introduced in order to normalize the slopes as the second performance will start at a slower AAT. This normalized comparison of the slopes of the first and second verse (blue and red respectively) clarifies the differences: the first's AAT declines from -290 *milliseconds* to -739 *milliseconds* or 449 *milliseconds* difference, while the second's AAT declines only from -862 *milliseconds* to -1048 *milliseconds* or 186 *milliseconds* difference. As Truslit predicts, on average, the amateur performers decelerated (relative to a fixed tempo baseline) for the first verse, and when performing the same material in the second verse, more or less, stop decelerating to keep the tempo somewhat constant through bars 12 and 13. Relatively speaking, they play the opening of the second verse faster than the opening of the first. As the above slope of the second verse of "O Holy Night" confirms, the rate of deceleration declines. Hypothetically, performance of the more complex sections of the piece, those sections with dyads, trains the amateur to perform such dyads, which they play more rapidly the second time.

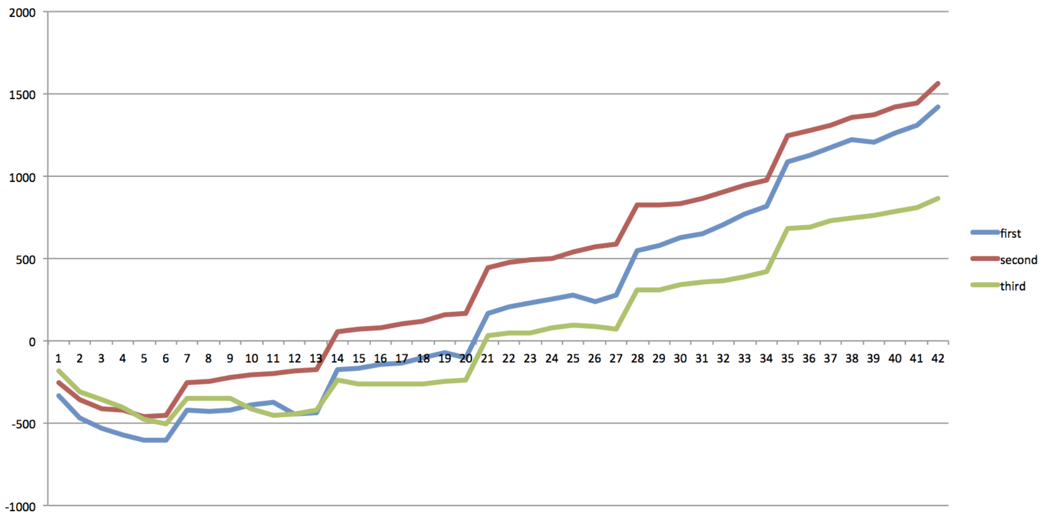

FIGURE 3.14: Slopes of repeat performances of "Twinkle".

Expanding upon the hypothesis that the repetition of more complex sections of pieces improves fluency, does repetition also inform an amateur's awareness of musical structure? To test this theory, we examined repeat performances of "Twinkle" and "Amazing Grace", another folk song. Truslit's belief that a performer will accelerate in a second

	accel at index 13	*AAT* at index 41
First performance	260.37	1423.55
Second performance	235.91	1563.26
Third performance	181.60	862.87

Table 3.2: *AAT* and acceleration in repeat performances of "Twinkle".

performance relative to a first is once again confirmed with these folk songs. An analysis of a third performance of each song, however, contradicts this trend: the *AAT* for the third performance is slower than the second, and remarkably in the case of "Twinkle", slower still than the first. Of the 26,867[5] performances of "Twinkle" analyzed, 510 performers were found who repeated a performance of "Twinkle" three or more times. Figure 3.14 graphs the slopes of this group's first, second, and third performance for "Twinkle". At note index 41 in "Twinkle", the first performance's *AAT* is $1,423.55$ milliseconds faster than the fixed tempo baseline at this index, as compared to the third performance, where the *AAT* at note index 41 is only 862.87 milliseconds faster, over a half-second slower than the first performance. In the case of "Amazing Grace", 481 of the 22,451 total performances repeated three or more times (see Table 3.1). As indicated by the graph (see Figure 3.15), a similar trend manifests in performances of this song.[6] The *AAT* of the third performance of "Amazing Grace" is slower than the second, and similar in tempo to the first, yet with more continuous arcs of accelerandi.

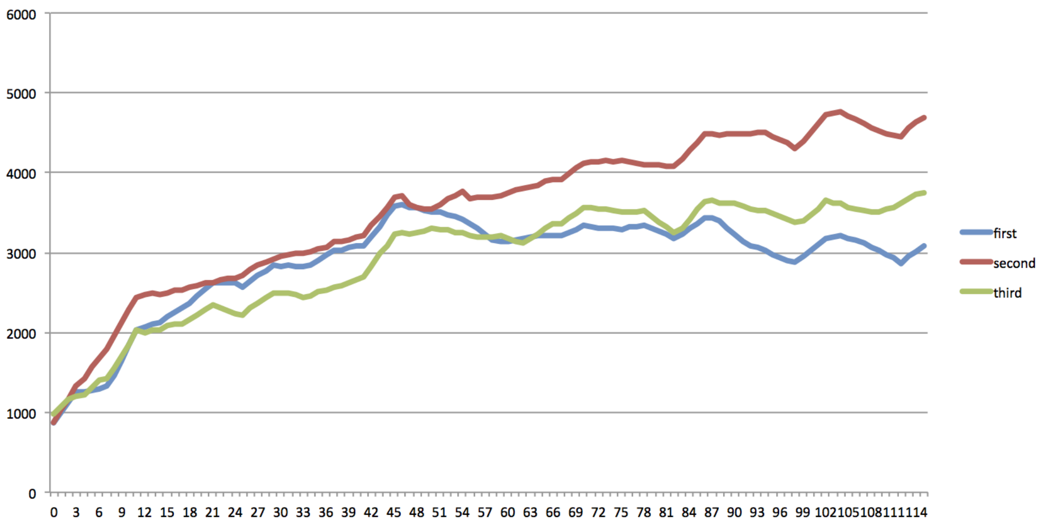

Figure 3.15: Slopes of repeat performances of "Amazing Grace".

[5] This number includes repeat performances. Previous calculations excluded all repeat performances and only included the first performance.

[6] As the "Amazing Grace" arrangement was all chords, the repeat analysis here used special software code to unpack these chords and thereby normalize all note values. A separate analysis that did not unpack any chords produced similar results.

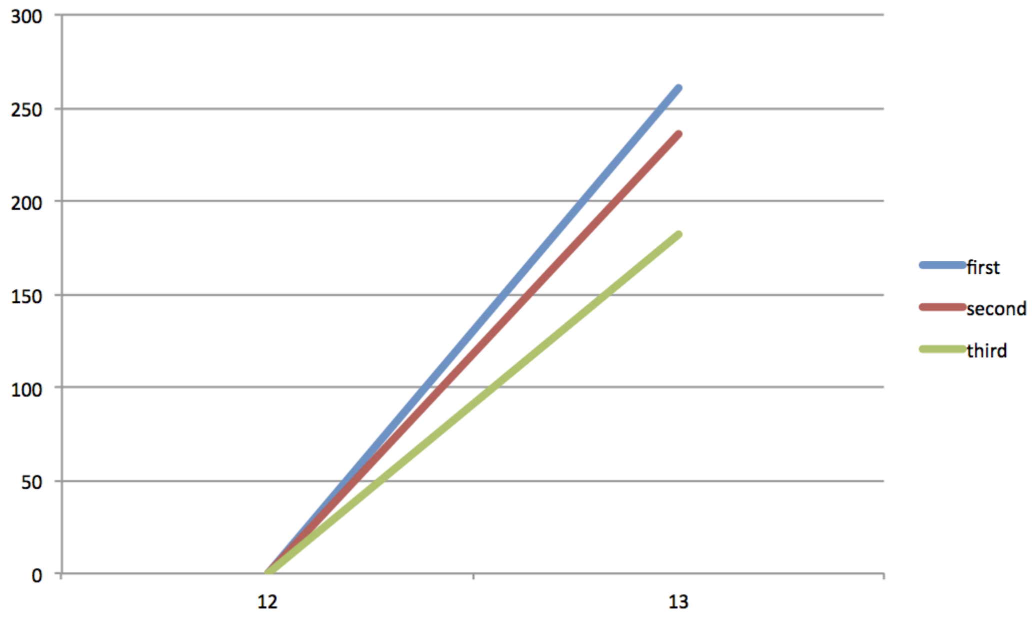

FIGURE 3.16: Normalized repeat slopes at phase boundary in "Twinkle".

While the tempi decreased on the third performance as compared to the first, the inflection points in the slopes where performers accelerate through each phrase boundary remain distinct. However, the rate of acceleration through these structural boundaries declines in absolute terms for each successive repeat performance. Table 3.2 shows the AAT at the end of each repeat performance, measured at note index 41, as well as the specific acceleration encountered at the second phrase boundary, or note index 13. This phrase boundary at note index 13 coincides with the end of the antecedent phrase and the return to the tonic — it is a significant structural moment in the piece. Figure 3.16 graphs the normalized AAT slope for each repeat performance at this phrase boundary.

Returning to our question about whether repetition informs an amateur's awareness of musical structure, our analysis of amateurs suggests that marking musical structure by decelerating at phrase boundaries in fact is a skill or an experientially developed behavior. However, while amateurs accelerate at phrase boundaries, performance repetition indeed seems to correspond to a reduction in phrase boundary acceleration, in some ways similar to the impact of repetition on the reduction of deceleration when performing dyads versus single notes. Yet, perhaps in part because of the mechanic, acceleration at phrase boundaries for amateurs seems endemic to the data. The unskilled performer does not appear to prioritize musical structure over basic technical proficiency.

Do these findings enlighten the view into intrinsic, non-learned behaviors? Possibly. Our study shows that amateurs do not prioritize structure. They accelerate when it is possible to do so. However, the analysis thus far has only considered the data in aggregate; in other words, the *average* amateur accelerates when it is possible to do so. Do the results also suggest that at some level basic skill is required in order to be expressive (i.e. to use agogics)? Perhaps, but once again this point must be qualified by the fact that the data has not been partitioned, thus far having been evaluated only in aggregate. It's possible, therefore, that some amateurs do not accelerate at phrase boundaries, and that most do, producing an average of acceleration.

Even so, the question of whether something is expressive, like the question of whether something is slow or fast, can be a relative one, not necessarily tethered to established standards of coherence. Put another way, expressive to whom? To an amateur or to an expert? Parsing what one hears is as much a skill as performing. For example, Sridharan et al studied the neural dynamics of event segmentation in music, and found that untrained listeners detected "course-grained transitions" — the boundaries between movements in baroque symphonic works (Sridharan et al.). Yet, in citing Knösche, they note that "these coarse-grained transitions are easily perceived by musically untrained listeners (non-musicians), unlike finer-grained transitions (such as "phrase transitions") that occur over shorter timescales of 1 s or less, which even musically trained listeners can find difficult to perceive" (Knösche et al.). And Barth, in *The Pianist as Orator*, writes that pianists often play notes of longer duration incorrectly in proportion to notes of shorter durations, yet they perceive that the note durations sound 'correct' (Barth).[7]

A canonical performance may produce coherence, albeit subjective coherence. But canonical should not be mistaken for intrinsic or universal. "Expression, to a great extent, is a matter of terms, and terms are anyone's. The meaning of 'God' may have a billion interpretations if there be that many souls in the world." (Ives) That is to say, if we temporarily suspend our knowledge (or, as Repp would have it, our learned behavior) and impose an atavistic standard in the place of a conventional standard, we might be able to glean some traces of universal behavior from this study.

3.5 identifying 'skill'

Emerson

The analysis thus far in this chapter represents only a cursory pass across the data. Many other theories could be tested. Several questions remain unanswered. For example, whether partitioning the performance data would reveal divergent treatment of agogics.

[7] "We "sound" our 2:1 ratios at roughly 1.75:1, but "hear" them as 2:1. [Footnote: In the extensive observations of professional and semi-professional performers by the psychologists Ingmar Bengtsson and Alf Gabrielsson, almost every performance of the notated 2:1 deviated markedly from this ratio (in most cases the longer note was shortened and the shorter note lengthened so that the ratio was much closer to 1.75:1), yet all of them "sounded" like 2:1.(Bengtsson and Gabrielsson)]" (Barth)

Having dismissed the notion that elucidating structure through the use of agogics is a universal behavior, we need to consider whether such elucidation, while learned, might nevertheless be intrinsic to certain regions (or cultures)? We'll pursue this question in Chapter 4. But first, we will explore whether it is possible to partition the amateur performance data according to the skills and experience of the performers. For while the performance data represents amateurs (see 2.2.3), certainly gradations of skill would be expected. Can such gradations be detected in the data? If they can, we might next ask whether the expertise we see represents performance skill, knowledge of the music, or even the influence of cultural norms. But for now, these questions must remain open.

Given that the use of ritardandi to communicate musical structure, including phrase boundaries, apparently reflects performance skill, the partitioning of performances by such a criterion could distinguish those amateurs possessing more skill (either expertise, knowledge, or both) from those with less. To test this hypothesis, the performances for a given song were divided into two groups. The first group, designated 'skilled', included all performances where the average note durations were above the AAT for a set of significant structural points in the piece. The expectation would be that performers who possess more skill might slow down at such points, for example to demarcate phrase boundaries. Only if the durations of all notes of a performance were above average for the entire set of significant structural points was the performance deemed 'skilled' — if the performer was slower at only one or two such points but not all, the performance was not considered skilled. Of course given that the preponderance of performance data demonstrates acceleration through such points, the test isn't filtering for the use of ritardandi in an absolute sense, but only relative to the average. The second group, designated 'unskilled', represents all performances that don't conform to this behavior, or by definition every other performance.

Performances of two songs, "Twinkle" and "Amazing Grace", were partitioned with this method in order to identify potentially skilled (i.e. knowledgable) groups. For "Twinkle" (see Figure 3.1), performance note durations at indexes 6, 13, 20, and 41 were analyzed. This set of note indexes corresponds to significant phrase boundaries in the piece. For a given performance, if all note durations in this set of were above the AAT, that performance was placed into the 'skilled' group. But if even one note duration of a given performance at any of these indexes was below the AAT, that performance was placed in the 'unskilled' group.

The results of the applied partitions are indicated in Table 3.3 and graphed in Figure 3.17. The table gives the number of performances, the percentage of urban performances, the average *Pearson correlation coefficient* (*A-Pcc* — see 3.2.1), and the AAT at index 13 for three groups: all performances, skilled, and unskilled. The *A-Pcc* calculation

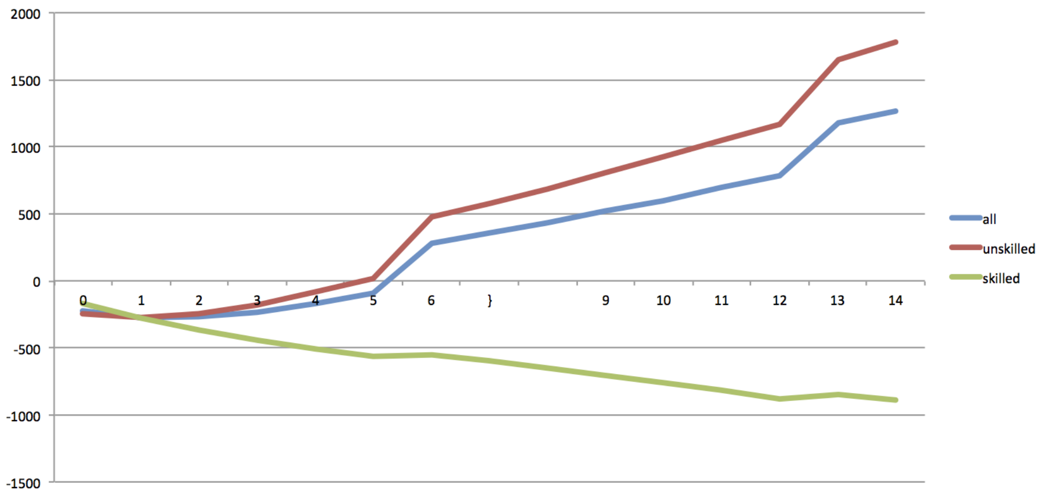

FIGURE 3.17: "Twinkle" slopes of skilled and unskilled performances.

	Performances	Urban %	*A-Pcc*	*AAT* at index 13
All	26867	58.4%	0.201	394.3
Skilled	5142	51.7%	0.667	29.4
Unskilled	21725	60.0%	0.137	480.6

TABLE 3.3: Skilled and unskilled partitions of "Twinkle" performances.

results suggest a weak affinity between performances in the unskilled group versus a strong affinity in the skilled group, more or less supporting the hypothesis. The slope of the unskilled group largely corresponds to the slope of all performances, albeit with additional acceleration at the phrase boundaries, which makes sense given the removal from this group of the set of performances that slow at all such boundaries. The *AAT* of the skilled group at the second phrase boundary (note index 13, the end of the antecedent and return to the tonic), was 29.4 ms as compared to the unskilled group of 480.6 ms, or acceleration relative to the *AAT*. The skilled group, by contrast and as expected, shows little to no acceleration relative to the *AAT* at each note index identified for the partition, the phrase boundaries. As an aside, the unskilled group also has a higher percentage of urban performances, or 60%, versus 51.7% for the skilled group. Two representative 'skilled' performances are included: one performed in Paddington, Australia and another in Safi, Morocco. If you listen to the full Australian performance, you will notice the performer accelerates rapidly through the second section where the arrangement returns with chords, yet perhaps expressively (?) the performer uses a pronounced ritardando to end the piece. The Moroccan performance, by contrast, uses a more consistent tempo throughout. In any case, relative to most other performances, these examples clearly demonstrate skill.

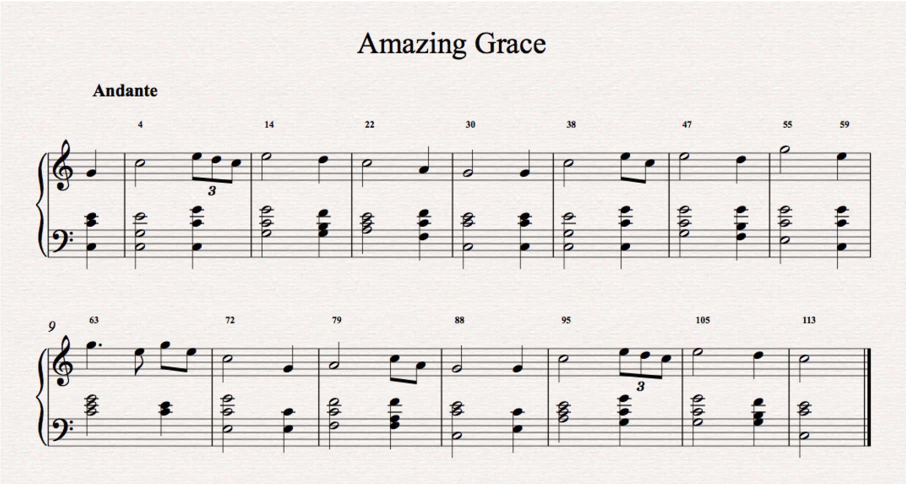

FIGURE 3.18: Note indexes for “Amazing Grace”.

Figure 3.18 gives the score and note indexes for the folk song “Amazing Grace”. These note indexes give the first note of a chord. Hence note index 33 is the last note of the chord beginning at index 30, or G, which would then be used to identify any ritarandi at this point in the piece. Applying the same partitioning technique to the song “Amazing Grace” yields similar results. Note indexes 33, 58, and 91 were used for the partitioning; they correspond to the primary phrase boundaries in the piece, specifically the return to the tonic from the subdominant/subtonic.

FIGURE 3.19: “Amazing Grace” slopes of skilled and unskilled performances.

The skilled group (see Table 3.4), like the skilled performers of “Twinkle”, has a much higher A-Pcc as compared to either the overall group or the unskilled group. The rate of acceleration through the phrase boundary at note index 33 is roughly half of the overall

	Performances	Urban %	*A-Pcc*	*AAT* index 33
All	23254	53.9%	0.293	−626.8
Skilled	6712	45.4%	0.890	1421.9
Unskilled	16542	57.4%	0.151	−1601.8

Table 3.4: Skilled and unskilled partitions of "Amazing Grace".

average acceleration at this point in the piece, and substantially lower than the unskilled group. Once again, the skilled group has a lower urban percentage of performances, in this case below 50 percent. Yet one significant difference stands out between the slopes of the skilled group and unskilled groups for this song (see Figure 3.19) in contrast to "Twinkle": on average the skilled performers play "Amazing Grace" faster than the unskilled performers, despite taking significantly longer at the phrase boundaries (which also coincide with notes of longer duration). Given the greater difficulty of playing chords in this piece, it's not surprising the skilled, while slower at phrase boundaries, exhibit greater proficiency in performing chords as well. More simply, this difference in tempo of skilled as compared to unskilled in both pieces is consistent with the hypothesis. One representative 'skilled' and two 'unskilled' performances of "Amazing Grace" are included: the identified skilled example was performed in McKinney, Texas, USA while the unskilled examples were performed in Sichuan, China and Holland, Michigan, USA.

Note that the 'unskilled' Chinese performer referenced in the above example rolls all chords. Moreover, this performance does not demonstrate any use of ritardandi at structural boundaries. By the definition of expertise as supported by research, this performance lacks skill and coherence. Yet from an aesthetic point of view, even absent such structural elucidation, the performance flows and may arguably be viewed as more pleasing to some (or many). Indeed, relative to the multitude of performances in the data, it could even be termed expressive. Such an observation, which seemingly contradicts our guidelines for detecting skill through the agogical enunciation of structural boundaries, in fact supports the broader thesis underlying this study of universal and intrinsic behaviors. We are reminded of the *us/them* dichotomy proposed by Agawu, as well as the Ives quotes that open this chapter. The identification of agogical enunciation of structure may indicate skill or universal expertise, or, it may indicate cultural bias in learned behavior, perhaps reflecting our own bias and standards. As opposed to identifying skill, the partitioning method described above may simply measure knowledge or even the "viscosity" of a cultural practice, the rate at which such a technique or value propagates across regions. Later chapters explore this alternative view into 'skill' or 'expertise' in more detail. But before considering concepts of cultural import or export and information viscosity, we will first further partition the performances again, this time seeking to analyze and correlate such behaviors to a second variable, namely ‘place’.

Chapter 4

Intrinsic behavior

Oh! virágok! Oh! ilatoskert!

4.1 outside/in

Canope

The search for the remote Transylvanian village continues. The previous chapter challenged theories of universal, and presumably innate behavior, at least in regards to agogics. This chapter focuses on local and intrinsic behaviors, which, by inference, suggest possible generalities.

Inspired by Bartók's methods of discovery and cataloging of culturally unique music features, we seek to correlate musical behaviors, specifically the treatment of time by amateur performers, with the geographic origins of these performances (Aarden and Huron). Whereas Bartók's methods could be characterized as an inside/out approach — that is, recording folk material at its source and then disseminating it — this study will adopt an outside/in approach, importing musical materials to a region, whereupon the performance interpretation of those materials within this region are analyzed and compared to that of other regions (see Section 1.4). We thus introduce a secondary variable — 'place' — to compare performances, and by assessing potential regional differences suggest some theoretical observations.

We first consider rather expansive definitions of regions, namely hemispheres or continents, and subsequently progressively reduce our scope to more localized regions such as sub-continents, countries, cities, and finally localized areas within a particular region. In addition to agogics and the associated agogical analysis to evaluate skill or

	Tempo	Skilled	*A-Pcc*	Urban
N. America	90.5%	31.1%	0.352	35.4%
Europe	93.4%	25.3%	0.327	41.0%
Australia	95.0%	33.5%	0.324	65.3%
S. America	101.0%	19.9%	0.309	73.7%
Africa	102.5%	14.3%	0.237	78.0%
Asia	105.6%	13.0%	0.213	74.7%

TABLE 4.1: "Amazing Grace" performance tempi by continent.

knowledge (see Section 3.5), we will also examine tempi. A useful baseline to evaluate the comparison of performances between regions is simply the aggregate average for all performances. A given regional value (e.g. tempi, skill, affinities, etc.), then, can be analyzed relative to that of all regions.

4.2 continents

Brouillards

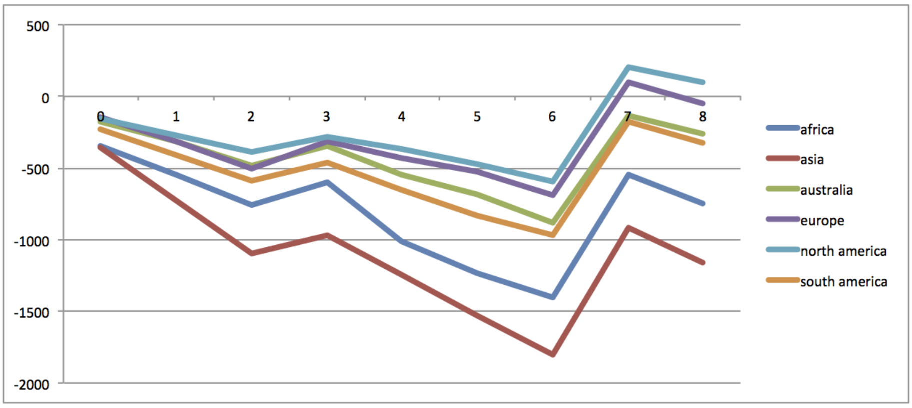

FIGURE 4.1: "Amazing Grace" performance slopes by continent.

Do amateur performance tempi vary by continent? Table 4.1 shows the tempi by continent for performances of "Amazing Grace". A high degree of variance exists in tempi across the continents. Performance tempi were fastest in North America, slowest in Asia, and average in South America.[1] There appears to be a relationship between skill and performance tempi, which makes sense given the chordal structure of this arrangement (see Section 3.4), in spite of the fact that we detect skill by looking for performances that have above average note durations at phrase boundaries (see Section 3.5). There

[1] Recall that Formula 2.6 gives the tempo as a percentage above or below the average of all performances for a given song. By this definition, a tempo of 95% would be faster than the average tempo by 5%.

also appears to be a relationship between the concentration of performances in non-urban regions (see Section 2.5) with both tempi and skill: non-urban regions, in general, demonstrate more expertise.

To gauge the level of similarity between performances within a region, the average *Pearson correlation coefficient*, or *A-Pcc* (see Appendix A.2), was also calculated for each continent. Once again, a relationship between tempi/skill and the affinity of performances within a region is apparent in the data: those regions demonstrating more skill have a stronger affinity between performances in the region. For example, the *A-Pcc* for North American performances is the strongest of all continents at 0.352; North America is also the region with the highest percentage of skilled performers. Figure 4.1 graphs the slopes for performances of "Amazing Grace" by continent. The graph confirms the differences in tempi by continent, with Asia the slowest and the most discontinuous in terms of acceleration and deceleration.

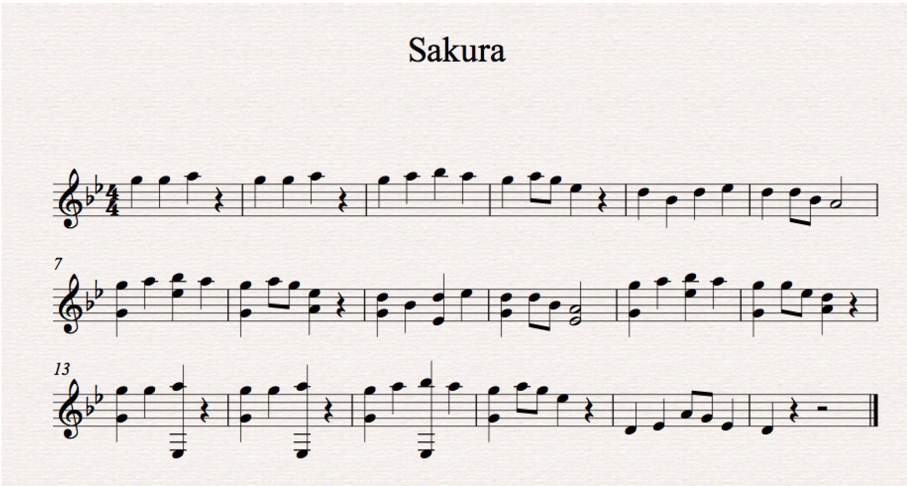

FIGURE 4.2: "Sakura, Sakura" score and note indexes.

An inter-continental comparison of performances of the traditional Japanese folk song "Sakura, Sakura" (see Figure 4.2) reveals significant differences in tempi as compared with those of "Amazing Grace". As shown in Table 4.2, the performance tempi across continents don't vary widely for "Sakura". Moreover, the *A-Pcc* for all continents suggests extremely strong affinities between performances. While the level of skill (again, as measured through agogics at phrase boundaries) ranges more than tempi or affinities — over 40% of Asian and North American performances are skilled, as compared with only 28.6% of South American performances — performances of "Sakura", unlike those of "Amazing Grace", show no prominent relationship between skill or knowledge and tempi. One likely inference to be drawn when contrasting these two folk songs: amateurs somewhat familiar with the music demonstrate more consistent interpretations. Given the strength of the affinities across all regions, a less obvious inference might be

	Tempo	Skilled	*A-Pcc*	Urban
Australia	98.0%	33.7%	0.808	68.4%
Africa	99.2%	26.9%	0.794	80.8%
S. America	99.9%	28.6%	0.801	70.9%
Asia	100.3%	43.6%	0.758	80.9%
Europe	100.6%	32.5%	0.824	40.6%
N. America	100.8%	45.0%	0.824	37.7%

Table 4.2: "Sakura, Sakura" performance tempi by continent.

	Tempo	Skilled	*A-Pcc*	Urban
Asia	95.3%	20.9%	0.268	76.5%
S. America	99.2%	19.0%	0.337	73.5%
Australia	102.7%	22.1%	0.353	59.3%
N. America	103.0%	25.4%	0.352	28.0%
Europe	103.1%	17.2%	0.330	42.9%
Africa	105.5%	21.6%	0.309	86.0%

Table 4.3: "Für Elise" performance tempi by continent.

that amateurs in a Western region choose to perform "Sakura" because they are already familiar with the piece. That is to say, as a user of "Magic Piano" can search for and choose which song to perform, the expertise and affinities demonstrated across the performances of songs may be self-selected, particularly for more obscure songs (see Section 2.2.2). The graph of the slopes of performances of "Sakura" by continent, like Table 4.2, demonstrate strong affinities for performance interpretation across regions (see Figure 4.3), supporting this theory. The performance interpretations for this song, too, are remarkably consistent across all macro-geographies. A representative skilled performance of "Sakura" from Kasugai, Japan is referenced.

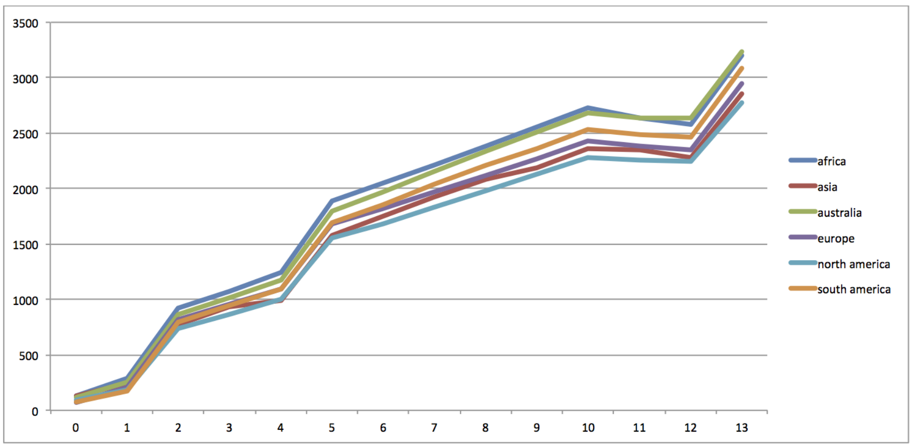

Figure 4.3: "Sakura, Sakura" slope by continent.

Table 4.3 provides a final example of performance tempi by continent, in this case for the classical piece "Für Elise" by Beethoven. Affinities between performances within a region, as measured by the *A-Pcc*, are low across all continents for this song. Asia, by far, demonstrates the fastest performances, and Africa the slowest. North America and Europe have similar tempi. Importantly, there are substantially more performances of "Für Elise" and "Amazing Grace" than there are of "Sakura" in the corpus (almost eight times more). This difference not only reduces the variance and statistical accuracy of analysis for "Sakura", but also buttresses the above theory of self-selection. For these reasons, more popular songs with a greater numbers of performances receive the most scrutiny in this study. Similarly, we often consider song genres as opposed to individual songs, noting that while a genre perhaps loses some specificity, it affords a greater number of performances and interpretations. In any case, tempi, in general, do vary by continent, and such variance seems related to performance affinities within a region as well as the percentage of performances that demonstrate skill or knowledge. Notwithstanding, while the categorization of performance tempi by continent offers some insights into regional behaviors, given the scale of such regions, these insights may be so general that they're of limited relevance. A progressively more granular regional analysis, therefore, follows.

4.3 north versus south

La puerta del vino

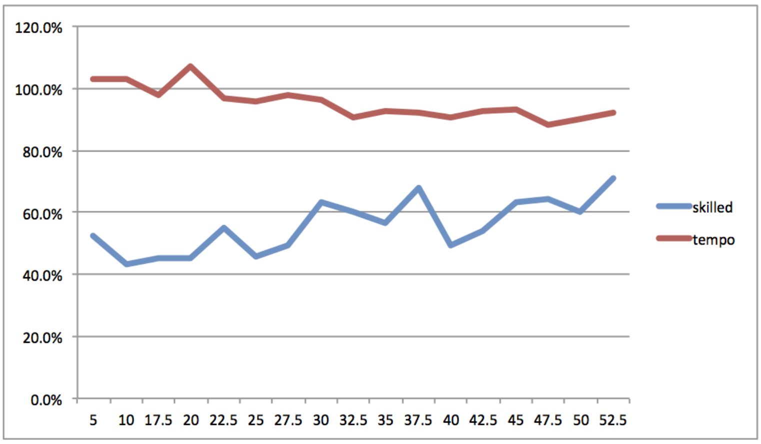

FIGURE 4.4: "Amazing Grace" tempi and skilled percentages by degree latitude, N. America.

Do the tempi and agogics of amateur performances vary by degree latitude? Depending on the region and the song performed, apparently so. Figure 4.4 graphs amateur performance tempi and skilled percentages for North America by degree latitude for the song "Amazing Grace". An apparent dependence manifests between degree latitude and tempo — in this case a *Pearson correlation coefficient* of -0.839[2] — suggesting that, for amateur performances of this song, the further north the performance, the more likely it will be faster than average, as indicated by performance tempi below 100%. The graph also shows a significant dependence between skilled performances and latitude/tempi, although slightly less pronounced. The *Pearson correlation coefficient* between skilled performance percentages and degree latitude is 0.707.

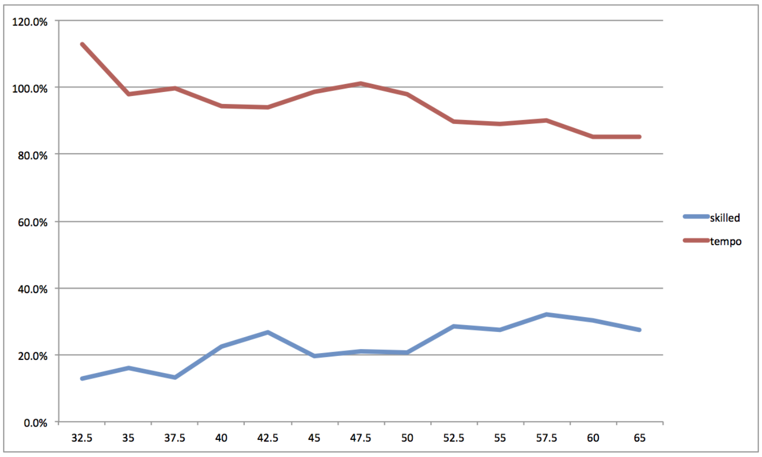

FIGURE 4.5: "Amazing Grace" tempi and skilled percentages by degree latitude, Europe.

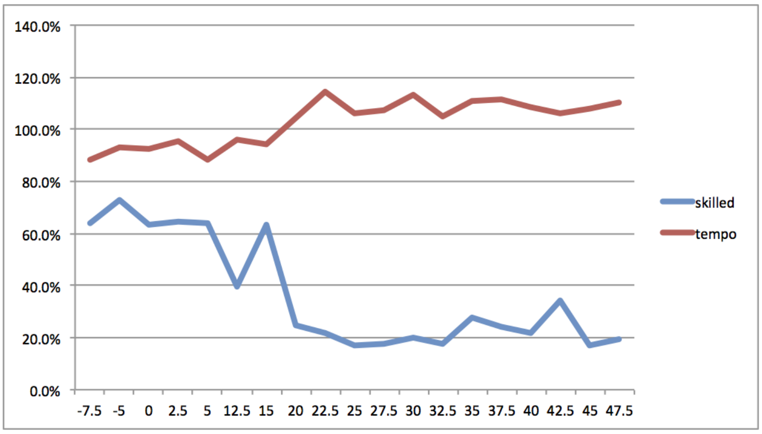

FIGURE 4.6: "Amazing Grace" tempi and skilled percentages by degree latitude, Asia.

The trend in Europe is almost identical to that in North America for performances of "Amazing Grace", as shown in Figure 4.5. The *Pearson correlation coefficient* between degree latitude and tempi is -0.797, strikingly similar to the dependence demonstrated

[2] Note that the correlation coefficient, in this case, is comparing the array of latitude values to the array of performance tempi, and is not to be confused with the method of calculating affinities of performances within a region through the average correlation coefficient.

	tempi correlation	skilled correlation
N. America	−0.839	0.707
Europe	−0.797	0.724
Asia	0.849	−0.853

TABLE 4.4: "Amazing Grace" tempi and skill correlations by latitude and continent.

in North America. Much the same, the link between demonstrated skill and degree latitude is also significant in Europe, with a *Pearson correlation coefficient* of 0.724. Asia also demonstrates a strong dependence between tempi and degree latitude (see Figure 4.6); however, in contrast to North America and Europe, the dependence is inverted: the northern performances are the slowest; the southern are the fastest. Likewise, a strong dependence between latitude and skill exists. This dependence is inverted as compared with North America and Europe for latitude, but congruent with North America and Europe for tempi and skill. The respective *Pearson correlation coefficients* for the relationship between degree latitude and tempi, and degree latitude and expertise are 0.849 and −0.853 for Asia. Table 4.4 summarizes the comparisons of "Amazing Grace" among North America, Europe and Asia.

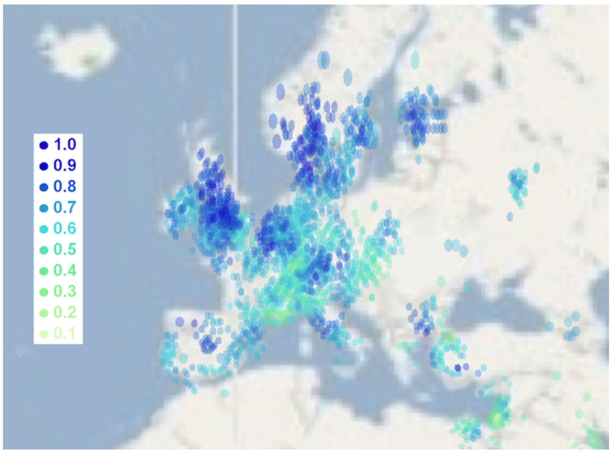

FIGURE 4.7: "Amazing Grace" tempi by region, Europe.

Figure 4.7 graphs tempi by region in Europe. Tempi were normalized to a scale of $0.0 \leq 1.0$, where unlike the convention used to measure relative tempi (see 2.6), 1.0 indicates the fastest tempo and 0.0 the slowest. The colored graphs show the contours of tempi changes throughout each continent. As indicated, tempi for the song are generally fastest in Scotland, Scandinavia, and Holland, and slowest in France, Central Europe, and Southern Europe. Figure 4.8 shows the same for Asia. According to the scale, for this song, the average tempi in regions throughout Asia are slower than those in Europe, with Malaysia, Indonesia, and the Philippines showing the fastest tempi.

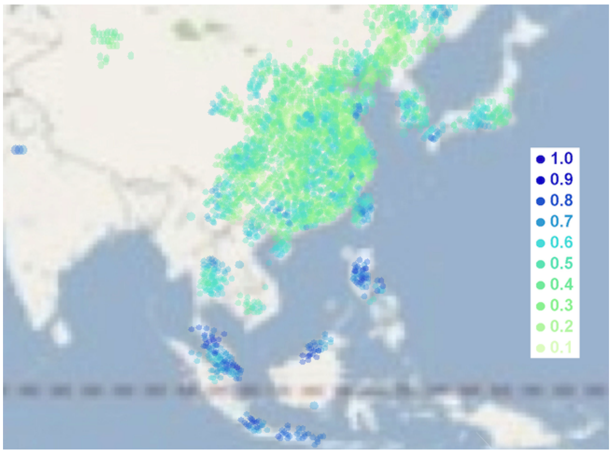

FIGURE 4.8: "Amazing Grace" tempi by region, Asia.

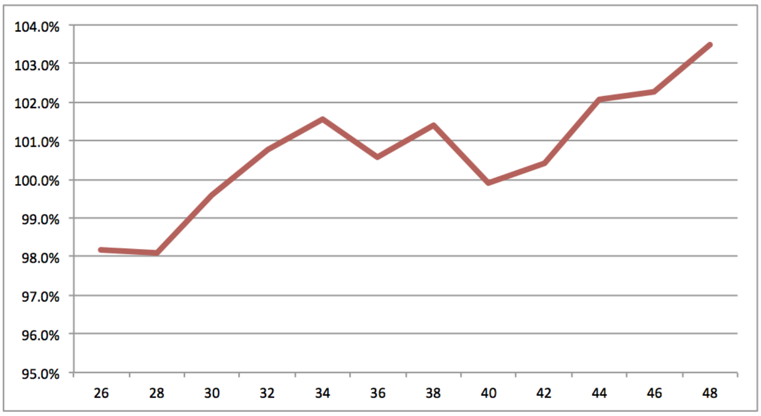

FIGURE 4.9: "Mario Brothers Theme" tempi by degree latitude, U.S.A.

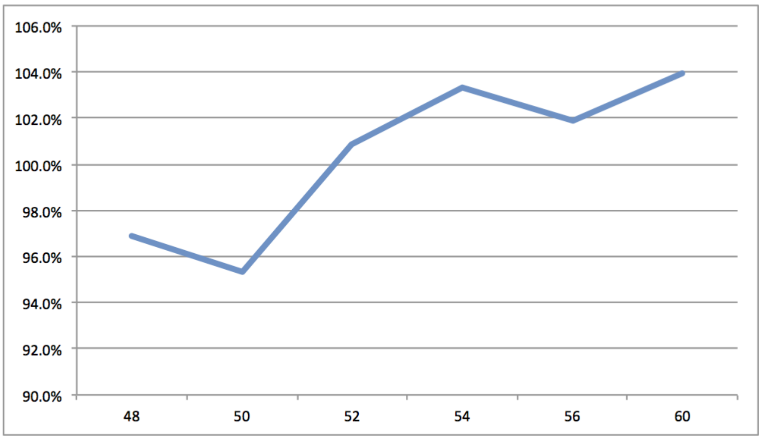

FIGURE 4.10: "Mario Brothers Theme" tempi by degree latitude, Europe.

"Twinkle" was also analyzed by degree latitude in these regions; similar dependencies were found to exist but are less pronounced.[3] However, other songs demonstrated strong dependencies between latitude and tempi in North America and Europe. For example, performances of the video game song, "Mario Brothers Theme", are graphed for U.S.A. and Europe, respectively, in Figures 4.9 and 4.10. In place of a score, a representative skilled performance of the song from Kempele, Finland is referenced. While the dependence between latitude and tempi remains evident for performances of this song in U.S.A. and Europe, in direct contrast to performances of "Amazing Grace" in this region, however, the tempi are fastest in the south, slowest in the north.

Aside from their distinct genres, an important difference between the songs relates to the arrangements: "Amazing Grace" is almost exclusively chordal, while in the "Mario Brothers Theme" single note strikes predominate.[4] In light of our definition of 'skill' in Section 3.5 (identifying 'skill'), we may theorize that the inverted correlation between latitude and tempi for these two songs stems from the more direct relationship of latitude to skill, as a skilled performer demonstrates mastery of chordal structures more rapidly, whereas a non-skilled performer will not respect note durations, particularly at phrase boundaries, and is more prone to accelerate through such musical structures. To put it simply, a skilled performer is more likely to play "Amazing Grace" faster than average and "Mario Brothers" slower than average. Several other possible theories exist that support the dependence between latitude and tempi or skill, including questions of why skill seems more prevalent in northern versus southern regions in the U.S.A. But short of this, it appears that degree latitude does indeed relate to performance interpretation, be this a manifestation of cultural norms, knowledge, skill, or possibly even intrinsic behaviors.

4.4 east versus west

Bruyéres

Do amateur performance tempi and agogics vary by degree longitude? As we saw in the analysis of the interaction between tempi and agogics with the degree latitude, depending on the region and the song performed, tempi and agogics also change with respect to the degree longitude. Figure 4.11 graphs amateur performance tempi and skilled percentages

[3] This is likely due to the arrangement of "Twinkle". The first section of single note strikes will be played faster by non-skilled and slower by skilled, who will pause at the phrase boundaries. The second section of the piece being largely chordal, the non-skilled will trade roles with the skilled and play this section slower, while the skilled will play it faster.

[4] In the arrangement of the "Mario Brothers Theme", the mechanic by default allows chords to be played with a single note strike as opposed to "Amazing Grace" where the chords are presented as multi-note strikes.

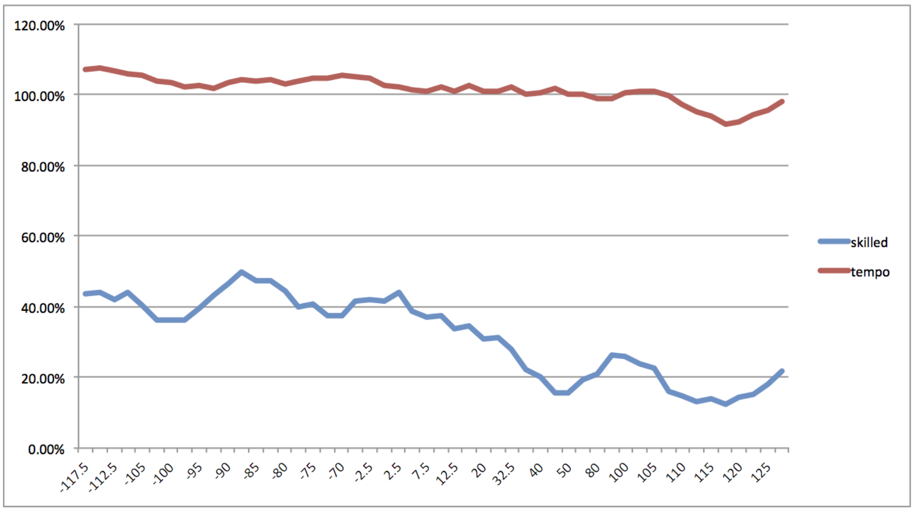

FIGURE 4.11: "Twinkle" tempi and skilled percentages by degree longitude, N. Hemisphere.

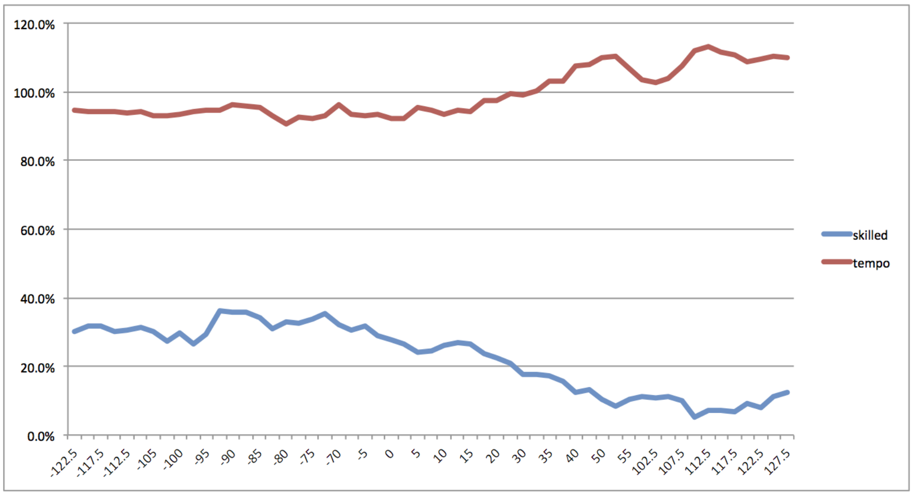

FIGURE 4.12: "Amazing Grace" tempi and skilled percentages by degree longitude, N. Hemisphere.

for the northern hemisphere by degree longitude for the song "Twinkle".[5] The tempi decline fairly consistently as the regions move from longitude −117.5° to 115°, where the tempi begin to climb. The inflection point at longitude 102° corresponds with the onset of significant population centers in China and Asia. Such regions demonstrate faster tempi and lower skills, as measured by the methods discussed in Section 3.5. The faster tempi, in this case, probably indicate aggregate acceleration through the notes of longer duration at the phrase boundaries of the song. Also note the inflection points at longitude −105° and −97.5°, where we see a decrease and subsequent increase

[5] A four-region moving average was used for some graphs and for the correlation analysis in this section. As each longitudinal region represented 2.5°, the moving average of four regions included 10° longitude. This moving average was used in Figures 4.11, 4.12 and 4.15.

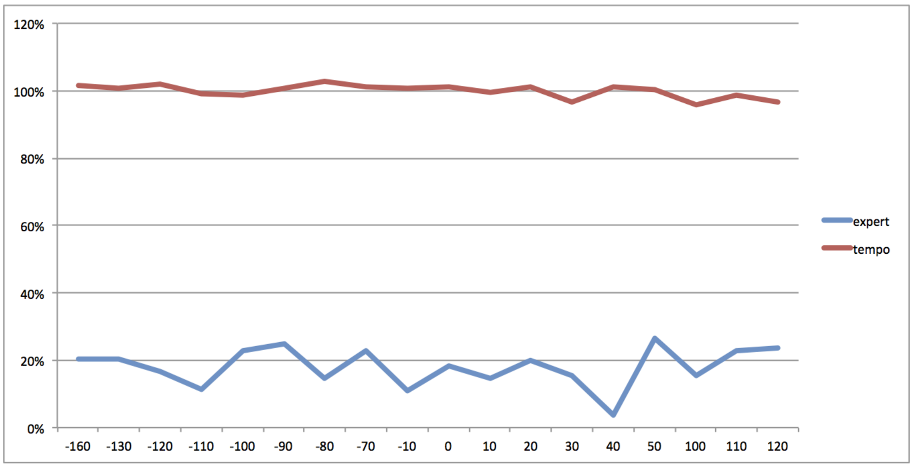

FIGURE 4.13: "Sakura" tempi and skilled percentages by degree longitude, N. Hemisphere.

in expertise and tempi; such inflection points correspond with the population centers in Mexico and Latin America, which when combined with the rest of North America in the longitudinal slice, apparently lower the tempi and expertise for those strands of longitude. Similarly, we notice the inflection point in skill at longitude 45°, the boundary between Europe and Asia.

A graph by degree longitude of amateur performance tempi and skilled assessment (in percent) for the northern hemisphere for the song "Amazing Grace" is consistent with that of "Twinkle" for skill, but inversely so with respect to tempi, which are fastest in North America, slowest in Asia. This inverse relationship between tempi and skill for "Amazing Grace" versus "Twinkle" can be interpreted as reflecting the more complex chordal structure of the arrangement. A skilled performer, as discussed in Section 3.4, will play complex structures more rapidly than a less skilled performer. Ironically, again, in this case the skilled performances are faster in spite of taking more time at phrase boundaries. A graph of the same for "Sakura" (see Figure 4.13) shows similar dips in skill at longitude −100° and 45°, however the inflection point and change in slope at the onset of signification population centers in China, roughly longitude 102°, are more modest. In addition, the tempi (as we saw in Section 4.2) do not vary greatly across degree longitude, perhaps due to the aforementioned reasons.

From the above, both latitudinal and longitudinal regions demonstrate a relationship (albeit sometimes an inverse relationship, depending on the arrangement) between tempi and skill, and similarly between skill and location (i.e. degree latitude or longitude). However, the relationship between tempi and latitude or longitude is not a correlation per se. Rather, the correlation is between skill and latitude or longitude, with tempi dependent on skill. Given the prominent inflection points that correspond to known cultural variations across such regions (for example at $-100°$ longitude, an inflection point that probably represents the population centers of Mexico and Latin America), the relationship between skill and degree latitude/longitude, it would seem, depends on variations in culture. In other words, while dependencies obviously exist between tempi/skill and region as defined by latitude/longitude, it's not clear those dependencies relate to east versus west or north versus south but rather the culture(s) identified by those regional slices.

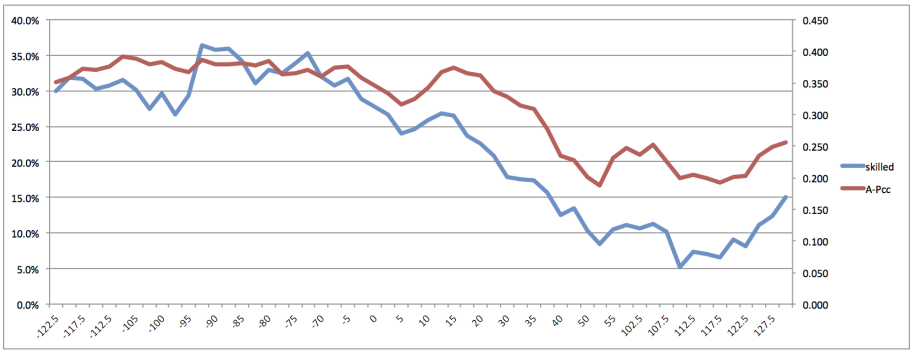

FIGURE 4.14: "Amazing Grace" skilled percentage vs *A-Pcc* by degree longitude, N. Hemisphere.

4.4.1 measuring cultural diversity

Assuming that the inflection points that show changes in tempi and skill in the longitudinal graph reflect differences in the homogeneity of culture for a given region, is it possible to detect such differences? For example, if tempi vary significantly in adjacent longitudinal regions, and such regions represent different combinations of culture, can the diversity or homogeneity of culture be characterized? This line of reasoning, of course, assumes that culture influences amateur performance interpretation, and that this perceived influence could be measured. To test this hypothesis, the similarity of performances within regions was calculated, again, using the *A-Pcc* to compare the coefficient between all performances in a region and then dividing by the number of comparisons, yielding a coefficient of likeness or affinity of performances within the region. Regions with limited or unique cultural influences, by this hypothesis, would produce a high coefficient (i.e. approaching 1.0), and those with diffuse cultures would produce a low coefficient — their performances would also be more diverse.

Taking input from the longitudinal graphs, performances in the region at longitude $-100°$ (a more heterogeneous cultural mix with a significant percentage of performances from Mexico in addition to the United States and Canada) were compared to those in a nearby region at longitude $-90°$ (a more homogeneous cultural mix with predominantly

	A-Pcc at $-90°$	*A-Pcc* at $-100°$
"Twinkle"	0.274	0.227
"Amazing Grace"	0.384	0.309
"Sakura"	0.815	0.737
"Für Elise"	0.369	0.326
"Mario Brothers Theme"	0.770	0.763

Table 4.5: Performance affinities at longitude $-90°$ and $-100°$, N. America.

American and Canadian performances) for North America. Table 4.5 reports the results of comparing the *A-Pcc* of these regions for several songs, and the results support the hypothesis: performance affinities are stronger in the more homogeneous region. The folk song "Amazing Grace" had the highest differences in affinities between the two regions (24%); the classical song, "Für Elise", less (13%); and the video game song, "Mario Brothers Theme", showed an insignificant difference (1%).

4.4.2 cultural diversity and skill

If the hypothesis holds, it would follow that regions demonstrating a higher percentage of skill, which in many cases implies knowledge, would also have a higher *A-Pcc*. Figure 4.14 graphs skill against *A-Pcc* by degree longitude for the performances of "Amazing Grace". The correlation between skill and *A-Pcc* is obviously strong (a Pearson coefficient of 0.893), thus providing more support for the hypothesis that cultural diversity has an impact on performance interpretations.

The integrated results of the analysis for latitude with those for longitude, which together examined approximately 100,000 amateur performances, show that while performance tempi vary by degree latitude or longitude, this fact appears to be a statistical dependence on skill or knowledge versus an actual correlation to location. Our more detailed study of the inflection points in these longitudinal graphs indicates that skill pertains less to degree longitude (or latitude) than on the level of cultural diversity in the given regional slice. Yet even so, recalling Agawu's critique of the Western scholar seeking to understand music from non-Western cultures, the definition of skill and its dependence on musical structure appears here strongest in Western regions, perhaps because we have defined skill as a functional of musical structure. Just the same, while the above findings motivate more questions about cultural bias and information viscosity (see Chapter 6), for now, once again a progressively more granular analysis of regions follows.

4.5 cities

La cathédrale engloutie

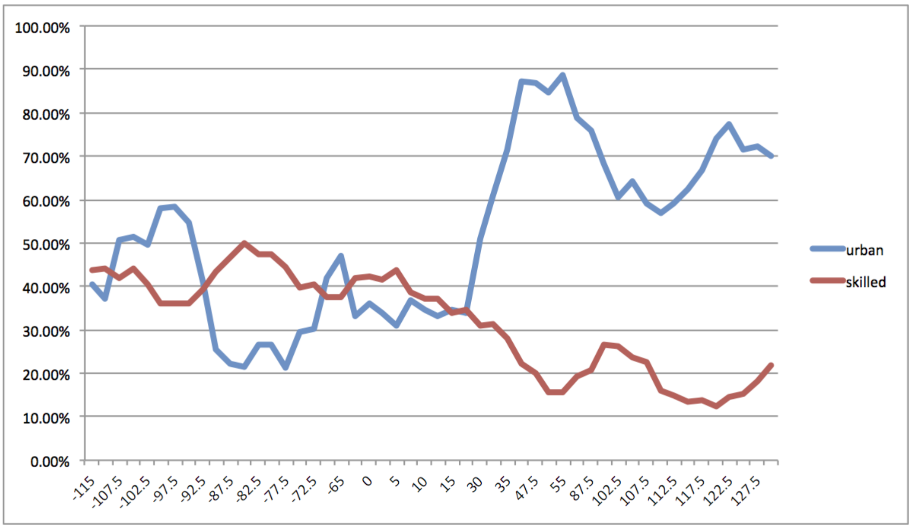

FIGURE 4.15: "Twinkle" skilled vs urban percentage by degree longitude, N. Hemisphere.

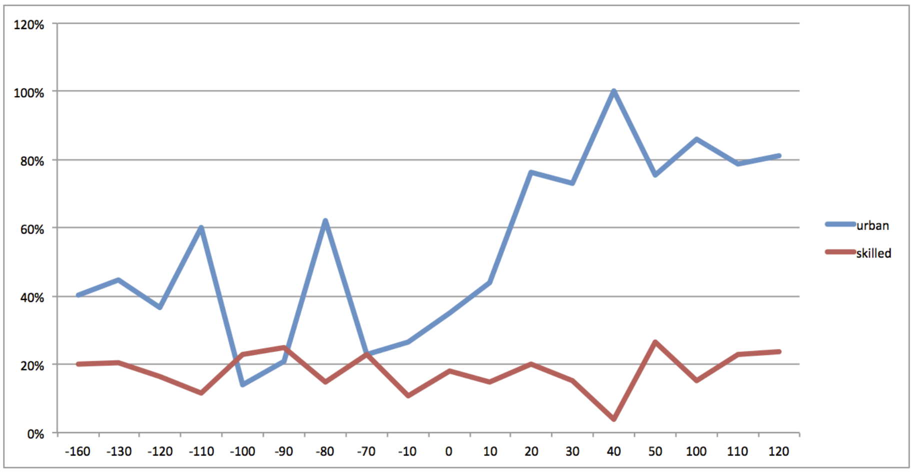

FIGURE 4.16: "Sakura" skilled vs urban percentage by degree longitude, N. Hemisphere.

A century ago, why did Bartók and Kodály leave the city and venture into remote Hungarian villages? Do urban regions lack culture? Returning to the hypothesis that the diversity of culture within a region corresponds to the level of knowledge as measured in agogical enunciation of structure, herein 'skill', we might, given trends in technology and immigration of the past century (see Section 1.1), expect urban regions to be more cultural diverse than rural regions and, correspondingly, possess less uniform and pronounced knowledge of the interpretation of a given song or genre. Taking the previous data from the longitudinal slices of the northern hemisphere, but instead comparing skill with the percentage of urban performances, we find a surprisingly strong negative correlation. Figure 4.15 shows that the higher the concentration of urban performances for a particular longitudinal region, the lower the level of skill or knowledge. The *Pearson correlation coefficient* (*Pcc*) for these two arrays of values (skill and urban percentages by longitude) is −.862. This relationship also appears in the "Sakaru" performance data as well (see Figure 4.16), although the dependence is less clear.[6]

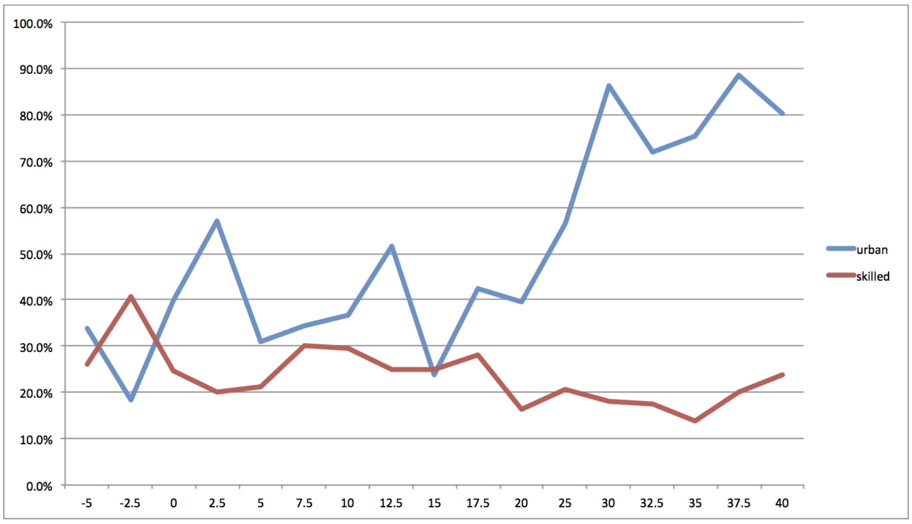

FIGURE 4.17: "Amazing Grace" skilled vs urban percentage by degree longitude, Europe

In order to remove the perturbation in the data from Asia, where in the case of "Twinkle" the spread between skill and urban concentration peaks, and where our marker for skill (time at phrase boundaries) may be less meaningful, a more granular view of skill versus urban concentration is given in Figure 4.17 for performances of "Amazing Grace" in Europe. The negative relationship between skill and urban concentration, via *Pearson*, still exists at −0.649,[7] providing further support for our hypothesis: rural regions with less cultural diversity have higher affinities across performances and an associated higher level of skill.

To increase the granularity, we shift from the latitudinal and longitudinal analysis, and instead divide the world into regions of approximately 120 square kilometers (or more precisely circular regions with a radius of 70 kilometers), analyzing every region with a statistically significant number of performances (see Section A.3). Distributing 5,000 performances of "Amazing Grace" across such regions produces 4,107 unique regions.

[6] Because only 3,059 performances in the corpus of "Sakura", as compared with over 20,000 each for "Twinkle" or "Amazing Grace", the regional slices of the data will be more crude.

[7] The graph of skill in Europe for "Amazing Grace" (see Figure 4.17) indicates an inflection point at around longitude 17.5°, or roughly the boundary between Western and Central Europe.

Partitioning the performances into two sub-groups, one comprising regions with no urban performances, the other comprising regions with only urban performances (see Section 2.2) yields 2,975 and 1,728 regions respectively. Correlation analysis is then run for the base group plus the two partitions, namely *all*, *rural*, and *urban*.

4.5.1 intraregional affinity

Thus far, the *A-Pcc* has been used to measure the degree of statistical congruence for a group performances (with respect to their use of time). A higher *A-Pcc* value for a group of performances, namely an average coefficient approaching 1.0, indicates a significant level of similarities between all performances in the group. An *A-Pcc* value approaching 0.0 indicates no similarities, and a value approaching -1.0 indicates, by contrast, differences in the group. Applying this definition to the set of performances within a particular region, we can characterize the *A-Pcc* for a group of performances defined by their region of origin as an *intraregional* affinity for the performances — that is, the degree to which the performances resemble one another *within* the region. Countries, cities, or 120 square kilometer areas could all constitute regions using this definition. For example, we could compare the intraregional affinity of performances of Western folk songs in the U.S.A. versus those in Ireland.

4.5.2 interregional affinity

The *interregional affinity*, in contrast to the intraregional affinity, indicate the statistical level of congruence of the use of time in performance interpretation *between* one region and a set of one or more regions. To compute the interregional affinity between two regions, A and B, the *Pearson correlation coefficient* is calculated between every performance in region A and every performance in region B. The correlation calculations are summed and then divided by the number of comparisons, yielding the interregional affinity (see Appendix A.2). Hence, the interregional affinity of a region is an indication of how similar performances in this region are to those in other regions. The interregional affinity, for example, could be the *A-Pcc* between every performance in Egypt and all performances in all other countries.

Returning to the more granular analysis of regional performances of "Amazing Grace", the *intraregional affinity* was computed across all performances within each region. The *average* intraregional affinity for all regions within each group was also calculated and is given in Table 4.6. This average intraregional affinity indicates the average level of performance congruence across all regions. The correlation between the intraregional affinity and skill was also calculated for each region.

	Regions	Average intraregional affinity	affinity/skill correlation
All	4,107	0.348	0.937
Rural	2,975	0.389	0.948
Urban	1,728	0.285	0.937

TABLE 4.6: Skilled percentage vs intraregional affinity affinity for “Amazing Grace”

The analysis of the partitioned groups confirms the hypothesis: *Rural regions, which theoretically lack the cultural diversity of urban regions, have much higher intraregional affinities across performances than do urban regions.*

Moreover, the higher the level of intraregional affinity (i.e. the higher the *A-Pcc* for performances within a region), the higher the percentage of skilled performances within the region. To confirm these correlations, the same analysis was run on “Amazing Grace” excluding Asian regions, noting the higher urban concentration and lower expertise in Asian regions, perhaps due to a lack of familiarity with the particular song. Excluding Asian performances yielded results consistent with the analysis that included Asia: higher average intraregional affinities. For example, excluding Asia, the rural average intraregional affinity was 0.488 versus an urban average intraregional affinity of 0.407. A regional analysis of “Sakura” that included Asia generated the same results, as did a regional analysis of other songs, including “Für Elise”. In other words, regardless of genre or macro-region, *urban areas consistently demonstrated lower intraregional affinities for performances as well as lower skill.* Given this, like Bartók and Kodály, we will leave the city and, after a brief tour or countries, examine rural regions in more detail.

4.6 countries

Des pas sur la neige

Despite the seemingly arbitrary nature of the diverse cultural composition of nation-states, this level of granularity, nonetheless, constitutes a regional definition for the comparative analysis of amateur performance interpretations.[8] Partitioning performances by country reveals variance by national borders. In particular, depending on the song and genre, the level of similarity between performances within a country – i.e. the intraregional affinity – varies considerably among different countries. Such differences relate both to the treatment of agogics and tempi.

[8] Consider, for example, changes over the past two centuries in the borders and population of Poland, or the immigration trends in Australia.

Initially, performances of "Twinkle" were partitioned by country. *Intraregional* and *interregional* affinities were calculated for each country. Table 4.7 gives these values for a few representative countries, and figure 4.18 graphs the slope of "Twinkle" for each. From the table, two inferences can be made. First, there is a relationship between intraregional affinity and interregional affinities. This makes some sense. For a region to have strong relationships with other regions, it must have identity. If instead the performance interpretations within a region are diverse, the intraregional affinity will be low or even negative, and hence such a region will lack an identity in terms of performance interpretation norms. Yet, the converse does not hold true: it is possible that a region with a high intraregional affinity has a cogent identity, yet that identity may be distinct, indicated by a lower interregional affinity. This theory is explored in more detail in the next Section (4.7). In general, however, we find that a high intraregional affinity corresponds to a high interregional affinity.

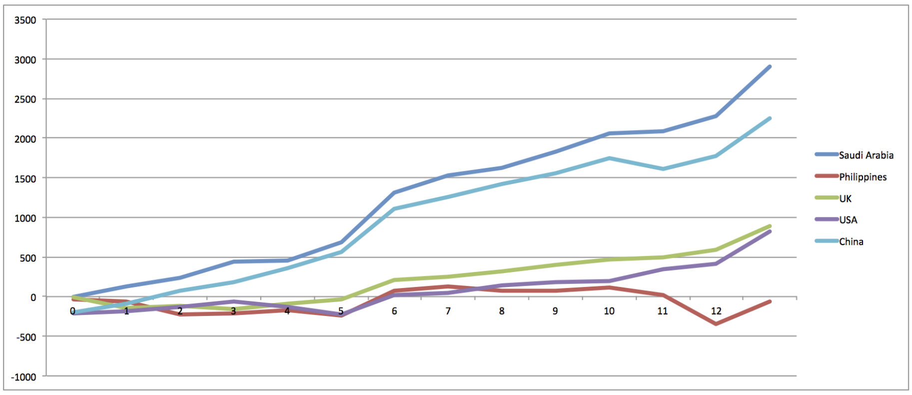

FIGURE 4.18: "Twinkle" slope by country.

Region	intraregional affinity	interregional affinity
Philippines	0.504	0.331
U.K.	0.410	0.304
U.S.A.	0.224	0.229
China	0.119	0.150
Saudi Arabia	0.075	0.133

TABLE 4.7: Intraregional affinities for "Twinkle" by country.

Second, there is a high degree of variance for the intraregional affinities among different countries — we find no wholesale migration to some global mean. The Philippines, for example, have an intraregional affinity (A-Pcc) of 0.504, almost seven times greater than that of Saudi Arabia for performances of "Twinkle". Agogics also differ substantially among the regions. For example, the slopes of the performances suggest that those in the Philippines decelerate prior to the phrase boundary, while those in the U.K. demonstrate no such inflection (see note indexes 5 and 12 on Figure 4.18), despite the fact that performances in the U.K. accelerate the least at the phrase boundary itself (i.e. are more 'skilled'). Performances in China and Saudi Arabia, by contrast, accelerate the most at the phrase boundaries (i.e. are the least 'skilled'). Once again we notice the relationship between our inferred skill and intraregional affinity. From such rough analysis, supported by the latitudinal and longitudinal analysis, we might expect the level of cultural diversity in each region might follow these values. That is to say, the lower intraregional values of Saudi Arabia and China may indicate greater cultural diversity (e.g. American oil workers?). Or, as an alternative theory, the absence of strong intraregional affinities in these two regions may reflect an absence of any knowledge of how to interpret such folk songs, largely of Western origin, and hence may signify a high standard of deviation for performance interpretation.

Region	intraregional affinity	interregional affinity
Philippines	0.567	0.420
U.K.	0.474	0.385
U.S.A.	0.455	0.378
China	0.109	0.176
Saudi Arabia	0.084	0.161

TABLE 4.8: Intraregional affinities for Western folk songs by country.

To test this alternative theory, namely the absence of interpretive knowledge for a given song within a region, performances of several folk songs were evaluated, in aggregate, by country (see Appendix C for a full listing of songs). Affinities, given in Table 4.8, were consistent with those for "Twinkle" (see Appendix D for intraregional rankings of all folk songs by country), although intraregional affinities in the U.S.A. increased somewhat as compared with those for the single folk song "Twinkle". Similarly, a comparison of several hundred thousand performances of several hundred songs of Western classical, Christmas, and video game music genres were all largely consistent with the analysis of folk songs by country for these regions, although in the classical genre, the values of China and Saudi Arabia were much higher in an absolute sense. Even for Western pop music, performances in China and Saudi Arabia remained in the bottom tier. Also of note, intraregional values in general were much higher for pop, video game, show, and Christmas music — there was much more diversity of interpretation, by contrast, for folk and classical genres. Yet for all genres of music, there is a high degree of variability of interpretation among different countries. Table 4.9 gives the intraregional values by music genre for the five countries.

Region	Folk	Classical	Games	Xmas	Pop
Philippines	0.567	0.730	0.771	0.776	0.733
U.K.	0.474	0.625	0.697	0.769	0.725
U.S.A.	0.455	0.674	0.709	0.803	0.759
China	0.109	0.346	0.568	0.379	0.493
Saudi Arabia	0.084	0.568	0.700	0.482	0.633

TABLE 4.9: Intraregional affinities by genre and country.

4.7 regions

Les collines d'Anacapri

Having thus far considered historical regional definitions such as continents and countries, here we depart from such definitions and adopt a significantly more granular approach to the study of 'place'. Performances were again distributed across approximate 120 square kilometer regions. For a given set of songs, we included each region that contained a minimum number of performances (see Appendix A.3), and for each such region, we calculated the tempi, skill, intraregional, and interregional affinities (see Appendix A.2).

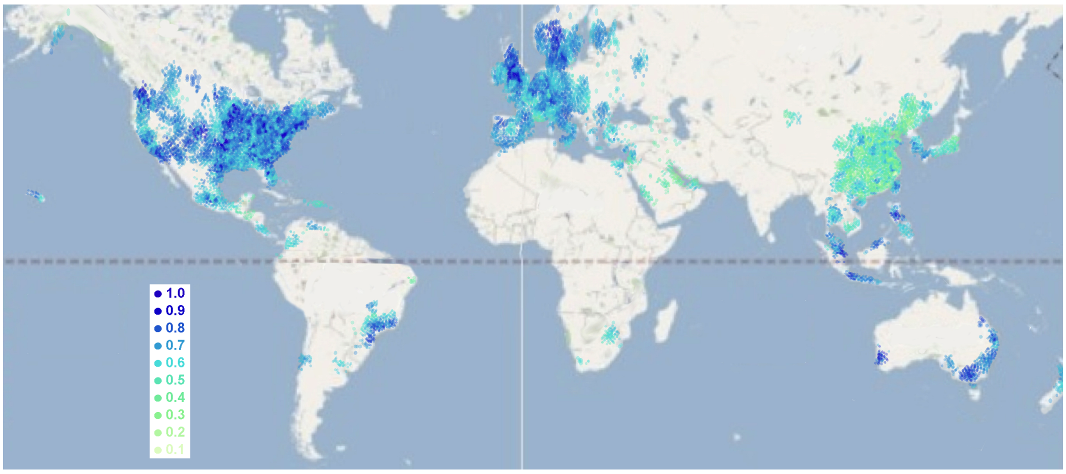

FIGURE 4.19: Regional tempi of Western folk music.

We began this more granular approach to region by evaluating tempi. Like we saw with continents, latitude, longitude, cities, and countries, tempi also vary by region.

Figure 4.19 graphs the regional tempi of roughly 123,000 performances of Western folk music songs (see Table C.6 for a listing of the folk songs). Again, note that the scale normalizes tempi, where 1.0 on the graph indicates the fastest relative tempi, 0.0 the slowest. The graph shows that tempi for Western folk performances are faster in Europe and North America than in Asia, but with a high degree of variability throughout world. For instance, in Europe, the fastest tempi for folk music emanate in Scotland, and the slowest in southern France. While the tempi vary by region, as noted earlier, this relation is not a direct correlation but instead more likely derives from the region's level of skill, knowledge, or perhaps even cultural norms, which in turn appear to depend on the region's level of cultural homogeneity. Hence we now turn to the question of whether cultural norms can indeed be identified with 'places' at this more granular level.

The search for the Transylvanian village continues, here via statistics. To identify regions that possess intrinsic cultural knowledge or behaviors, regions were sorted and ranked by intraregional and interregional affinities. Two cases were considered: first, those regions with a discrepancy between their intraregional and interregional ranks, and second those regions with both a high intraregional and interregional rank (i.e., no discrepancy between the ranks). The hypothesis for the first group: a strong intraregional rank would indicate identity, which combined with a weak interregional rank might designate a region with distinctive intrinsic behaviors (i.e. with behaviors that are not shared with other regions). The second group (high intraregional and interregional ranks) might identify regions with intrinsic behaviors that are shared by other regions, and hence a source of broader cultural norms or even universal behaviors. Only regions with high intraregional values were evaluated in depth, as regions lacking such strong internal affinities across performances, again, would lack an identity, rendering generalizations about performance practices in this region or comparisons to other regions impractical.

This statistical method of searching for intrinsic behaviors was applied to performances of "Amazing Grace". 28,917 performances produced 12,840 unique regions, of which 11,643 had the requisite minimum number of performances. Figure 4.20 graphs the intraregional affinities of these 11,643 regions. Noting the origin of the folk song, a Christian hymn written in Buckinghamshire, England, regions with the strongest cultural affinities to England predominate. Specifically, and in contrast with the graph of all performances (Figure 4.21), the leading intraregional affinities for this folk song are concentrated in the U.S.A., the U.K., and Northern Europe.

Of all regions, *Harper, Kansas, U.S.A.* (GPS 37.292892,-98.0) was the origin of the region with the highest intraregional affinity: 0.857. *Yorkshire, England* (GPS 53.646446,-1.3876276), *Penicuik, Scotland* (GPS 55.816986,-3.1830127), and *Vøringsfoss, Norway*

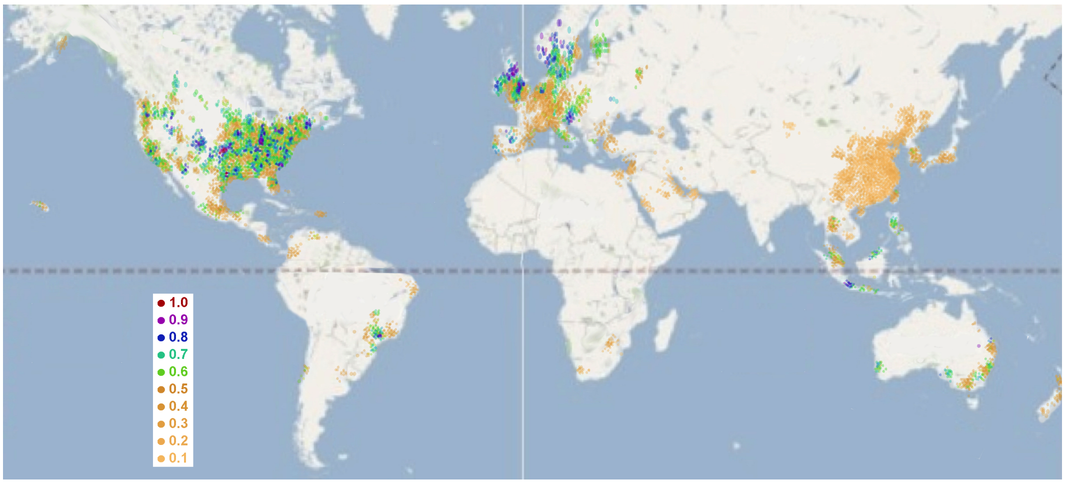

FIGURE 4.20: Intraregional affinities for "Amazing Grace".

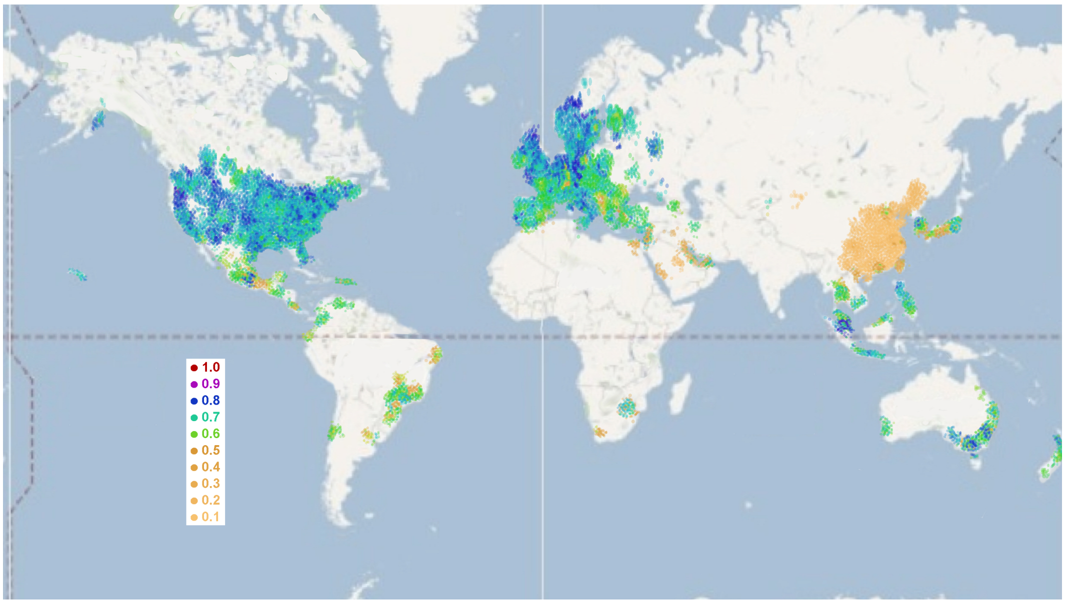

FIGURE 4.21: Intraregional affinities for Western folk music.

(GPS 60.207108,7.5) also all had high intraregional values, above 0.780. The average intraregional affinity for all 11,643 regions was 0.363. Near the middle of the spectrum we find *Casimiro de Abreu, Brazil* (GPS -22.5,-41.292892) with an intraregional affinity of 0.410, and at the low end of the spectrum, *Jiangsu, China* (GPS 32.5,119.79289) with a value of 0.096. Figure 4.22 graphs the performance slopes for these regions.

Importantly, the slope gives the average aggregate tempo. It represent neither a standard of deviation, nor the degree of similarity across a set of performances in the case of correlation analysis. Yet for the regions with high intraregional affinities, we see a high degree of congruence of slopes, with some basic deviation on the treatment of the first strong beat and the triplet in the second bar of the song (see Figure 3.18 for

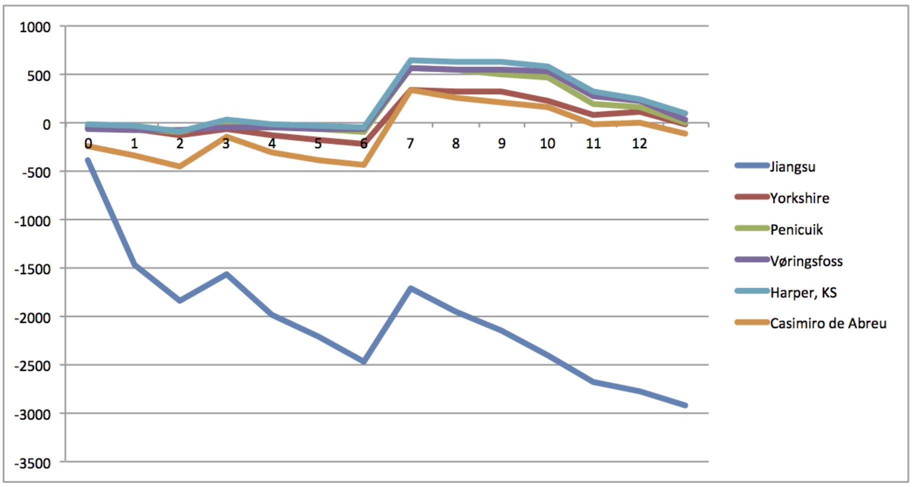

FIGURE 4.22: "Amazing Grace" regional slopes.

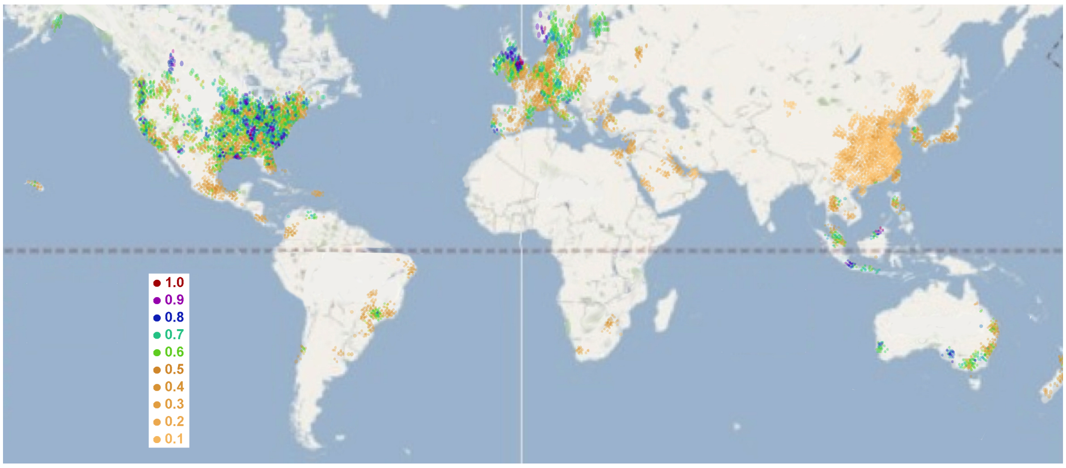

FIGURE 4.23: Skilled regions of "Amazing Grace".

the "Amazing Grace" score). The *Casimiro de Abreu* slope shows similar congruence but modestly departs from this set at note index 2. The *Jiangsu* slope is clearly not congruent with the slopes of the other regions. While the normative behavior for the performances (in aggregate) shows acceleration through the notes of the chords, by contrast a significant number of performers in the *Jiangsu* region are decelerating notes within the chords — they are in fact *rolling* the chords. Hence the standard of deviation across performances for this region is high, and the resulting intraregional affinity is low. Indeed, for much of the intraregional analysis for this corpus of data, the Chinese performances are anomalous. Quite possibly, this is a case of *them* versus *us* (Agawu).

Regional analysis of agogic use indicates that those regions with strong intraregional affinities have a much greater propensity to *ritard* (relative to the *AAT*) at phrase

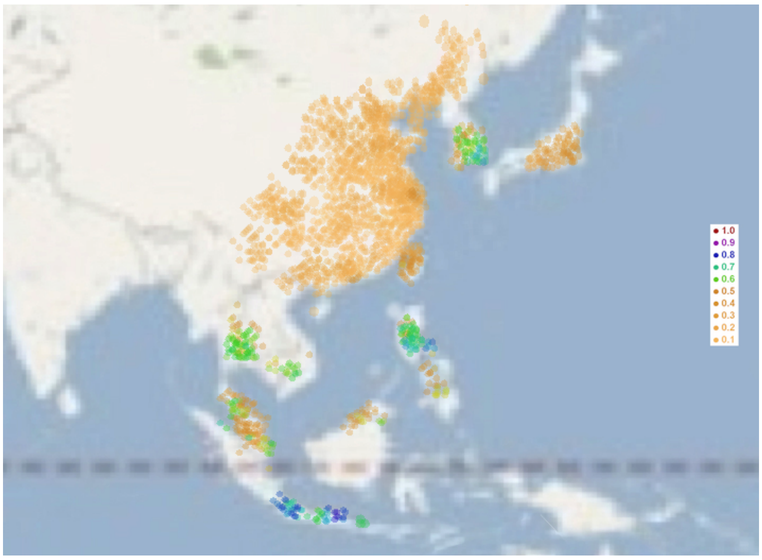

FIGURE 4.24: Intraregional affinities of “Für Elise” in Asia.

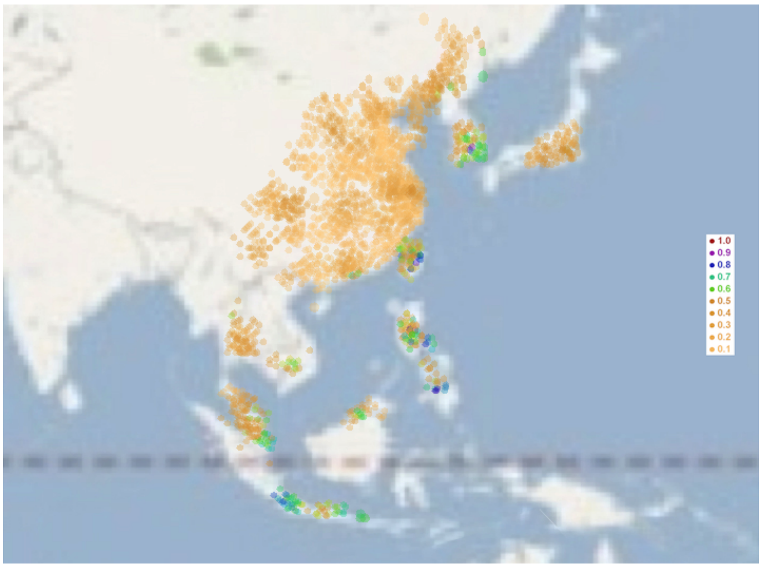

FIGURE 4.25: Skilled regions for “Für Elise” in Asia.

boundaries. Indeed, in performances of “Amazing Grace”, the correlation between intraregional affinity and skill (as measured by agogical emphasis of phrase boundaries, see Section 3.5) is significant, with a *Pcc* of 0.936. Figure 4.23 graphs regional skill. The relationship between intraregional affinities and skill becomes evident when we compare Figure 4.20 and Figure 4.23. The same correlation appears in analyses of other songs, including “Twinkle” and “Für Elise”. Figure 4.24 graphs intraregional affinity of performances of “Für Elise” in Asia, while Figure 4.25 graphs the *skilled* regions of performances (the scale for this graph has been normalized). The correlation between the two graphs is apparent. Regions that demonstrate higher intraregional affinities also are likely to exhibit the use of ritardandi at phrase boundaries, so much so that such behavior seems *intrinsic* to such regions for a given song.

If, as discussed earlier, the use of ritardandi at phrase boundaries is a learned practice (see Section 3.3), and by definition therefore not a universal behavior, then it is possible this practice is developed as skill through repetition. Or, it may be that the practice is intrinsic to a set of regions for a set of songs. In other words, the learned practice may be derived from cultural norms prevalent in such regions. Recall that this study, thus far, has demonstrated methods that can detect the use of such agogics to demarcate structure, and has shown, from a normative standpoint, that agogical enunciation of structure is not a universal behavior, at least not amongst the amateur performers surveyed in this research. Moreover, the prevalence of such agogic use, as articulated above, appears to depend (among other factors) on the level of cultural homogeneity within a region. And through this more granular view, those regions with higher concentrations of skill that also demonstrate higher intraregional affinities across performances can be distinguished from other regions. The confluence of these factors — the correlation between cultural homogeneity, intraregional affinity, and concentration of skill — therefore suggests cultural norms may be intrinsic to such identified regions.

To summarize, *while the use of agogics to demarcate structure is neither natural nor universal (i.e. it is a learned behavior), for a given set of songs, it appears to be an intrinsic behavior for a subset of regions.*

Given the potential of intraregional affinities to help us identify regions with intrinsic cultural norms, several songs were analyzed and graphed, here in Europe to simplify the regional comparisons.

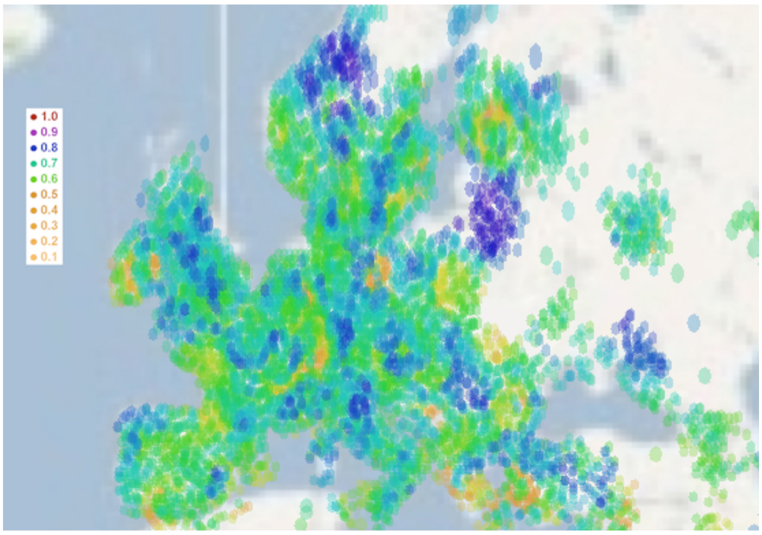

FIGURE 4.26: Intraregional affinities in Europe for Western classical genre.

- Figure 4.26 graphs intraregional affinities for 300,000 performances of approximately 127 Western classical songs (see Appendix C.1 for a full listing). From the graph, we see the strongest intraregional affinities in Northern Scandinavia and the Baltic countries.

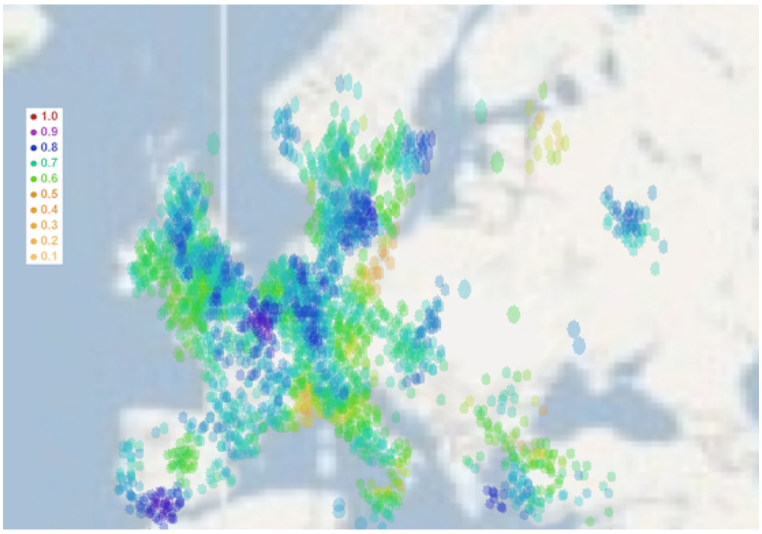

FIGURE 4.27: Intraregional affinities in Europe for "Moonlight Sonata".

- Figure 4.27 graphs those affinities for 20,000 performances of the "Moonlight Sonata" (Beethoven's Opus 27 No. 2 in C-sharp minor, first movement). From the graph, the strongest intraregional affinities are in Belgium, Central Germany, and Denmark.

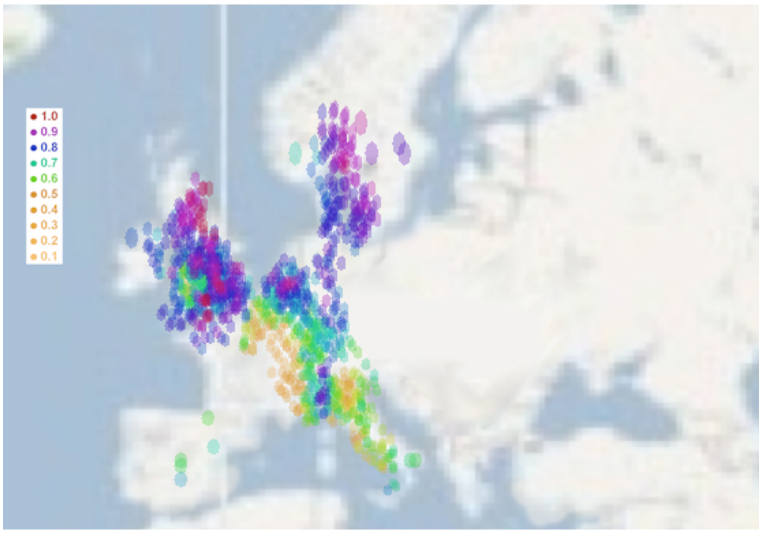

FIGURE 4.28: Intraregional affinities in Europe for "O Holy Night".

- Figure 4.28 shows intraregional affinities for "O Holy Night", a Western Christmas Carol. The strongest intraregional affinities appear within Scotland and Southern Scandinavia.

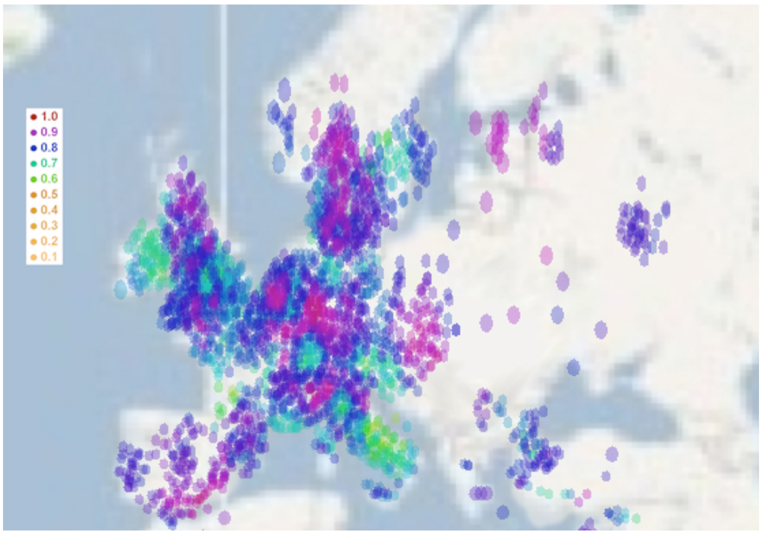

FIGURE 4.29: Intraregional affinities in Europe for Chopin pieces.

- Figure 4.29 shows intraregional affinities for 32,000 performances of fourteen classical pieces by Chopin, including several preludes in addition to a few etudes and waltzes. Central Germany, Poland, Southern Scandinavia, and Southern Spain show high intraregional affinities for these performances.

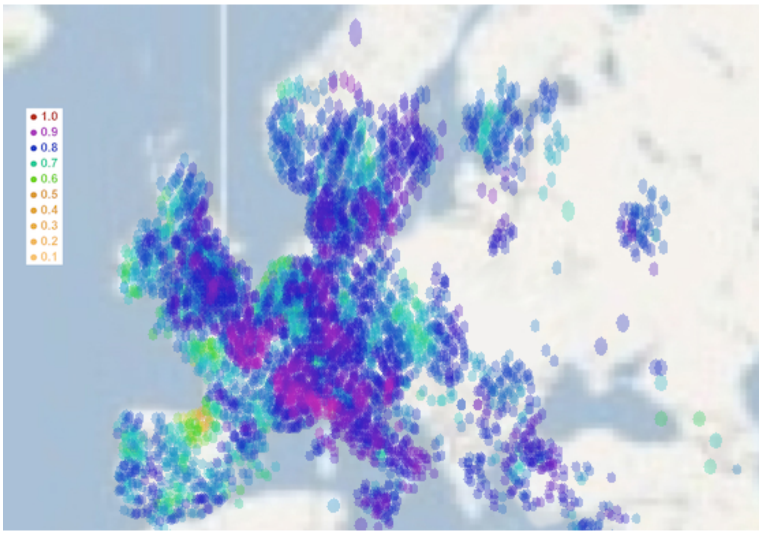

FIGURE 4.30: Intraregional affinities in Europe for Bach pieces.

- Figure 4.30 shows intraregional affinities for 55,000 performances of twenty-four classical pieces by J.S. Bach, including selections from the *WTC*, all movements

> from Partita No. 1 in B-flat major, two minuets, three inventions, "Air on the G String" (via August Wilhelmj), "Jesu Joy of Mans Desiring" (from *Herz und Mund und Tat und Leben*) and "Sheep May Safely Graze" (from Cantata No. 208). In some contrast with the intraregional affinities for Chopin performances, the strongest affinities for Bach performances in Europe are in North/Central France, Southern France/Northern Italy, with significant regions throughout Germany, Switzerland, and Denmark.

The preponderance of data indicates, by inference, that indeed cultural norms can be distinguished at a regional level through statistical analysis. Having confirmed that intrinsic behaviors correlate with the use of agogics to demarcate structure, several areas, of course, remain to be explored, including a more detailed study of agogical differences between regions that demonstrate such intrinsic behaviors.

In addition to agogics and tempi, other intrinsic behaviors can be analyzed and seem to be evident in the data. For example, deeper scrutiny of the treatment of the triplet gesture in the second bar of "Amazing Grace" reveals another such behavior that is non-normative and unique to a subset of regions. Figure 4.31 graphs the slopes of those regions for performances of "Amazing Grace". Virtually all performances, in aggregate, decelerate through the triplet, in particular during its second note, note index 12 (see Figure 3.19). That is to say, these regions perform the gesture as a triplet. The regional performances of *Vøringsfoss*, *Harper*, *Penicuik* and even *Jiangsu* conform to this normative behavior for this song. Performances in the *Yorkshire* and *Casimiro de Abreu* regions, however, accelerate through this second note of the triplet — they are not playing a triplet and instead treat the last note as a grace note to the neat beat.

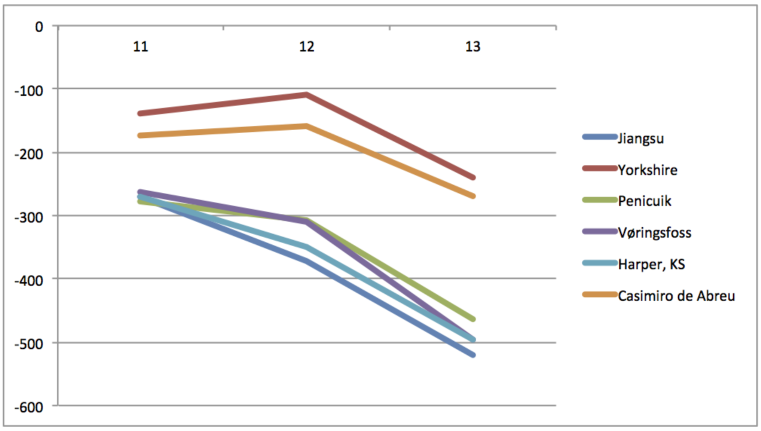

Figure 4.31: "Amazing Grace" triplet slope by region.

In any case, the theories of Aarden and Huron have born fruit (Aarden and Huron). The corpus of amateur performance data has yielded several potential sources of intrinsic behavior. Through the application (and inversion) of Bartók's own model — our *outside/in* approach — we have shown that performance interpretations do vary by 'place', and that such variation correlates to cultural homogeneity and skill. We infer agogical 'skill' to reflect a learned behavior (comprising practice and rehearsal) and/or cultural norms. In the case of the latter, the question remains of how such norms have been defined for particular songs or genres. For example, is the norm indigenous? Or instead, was a particular interpretation of a song simply imported into a particular region, perhaps via the radio or YouTube or by a 'rock-star virtuouso' (perhaps even from a nineteenth-century tour by Franz Liszt)? And if imported, did an indigenous cultural norm further influence it, the extent of its assimilation, or even whether it might be accessed? This line of reasoning suggests that the interpretative knowledge within a region is a function of information viscosity, the rate at which information flows into a region from abroad, which is dependent in part on cultural norms intrinsic to that region. We'll examine this in more detail in Chapters 6 and 7. But in order to frame such questions, we will first attempt to correlate identified behaviors across disparate regions.

Chapter 5

Regional affinities

Nézd hogy derül már a váram

5.1 theory

Meeresstille und glückliche Fahrt

The previous chapter attempted to correlate intrinsic musical behaviors with 'place', starting with continents and eventually considering 120 square kilometer regions. Performance interpretations, specifically tempi, agogics, and their associated impact on intraregional affinities, do vary by region. For example, the analysis of performances of the Western Christmas song "O Holy Night" within Europe revealed strong intraregional affinities in Scotland and Scandinavia (see Figure 4.28). Such strong intraregional affinities suggest a high degree of similarity across performances within these regions, which in turn correlates to both a higher percentage of skill or knowledge within the region (as measured through the agogical enunciation of phrase boundaries, see Section 3.5) as well as a greater degree of cultural homogeneity in the region (see Section 4.4.2). If these regions, which demonstrate strong intraregional affinities, indeed possess unique or even intrinsic behaviors relating to performance interpretation, is it possible to statistically search the globe for regions that bear similar intrinsic behaviors?

To frame the hypothesis, suppose the performances of Western folk songs in a particular region in southern Italy have a high level of similarity and perhaps unique performance features. What if the performances of this music in this region were compared to the performances in all other regions across the globe? Would regions with similar intrinsic behaviors manifest in the search? Or would all performances in general move to a broad mean, negating any value of such a statistical comparison?

To explore the hypothesis, three case studies were developed. The first two case studies used the corpus of amateur folk song performances (see Appendix C.6), or approximately 123,000 performances. The third case study used roughly 25,000 performances from the corpus of Western Christmas music (see Appendix C.3). For all the performances of these two genres, we compared affinities between performances in a given region to all other regions (the *interregional* affinity), once again using the *Average Pearson correlation coefficient* or *A-Pcc* (see Appendix A.2).

Using the statistical search techniques outlined in Chapter 4 (see Section 4.7) to identify regions with a strong identity, we first ran global comparisons of regions with high intraregional affinities. The search identified several regions with significant intraregional affinities for the Western folk songs, including performances near Askam-in-Furness, England (GPS 54.183014, -3.1830127) with an intraregional affinity of 0.912, Ashland, Oregon, U.S.A. (GPS 42.183014, -122.683014) with an intraregional affinity of 0.882, and finally Corigliano Calabro, Italy (GPS 39.707108, 16.5) which had an intraregional affinity of 0.878. Interregional affinities between these three regions and Europe and Asia serve as the first two case studies for this chapter. For the third case study, which analyzed global regional affinities of Western Christmas music, performances within the Kemnay, Scotland region (GPS 57.207108, -2.5) were compared with those from Australia.

5.2 case study: Europe

Piano Trio No. 1 in D minor Op. 49

Interregional affinities for Western folk performances were calculated between regions throughout Europe and the Askam-in-Furness, Ashland, and Corigliano regions. The results are graphed below.

- Figure 5.1 shows the interregional performance affinities between the Askam-in-Furness, England region and those of the rest of Europe. Most of Scotland, Norway, Austria, Slovenia, and Italy demonstrate high intraregional coefficients with Askam (above 0.80).

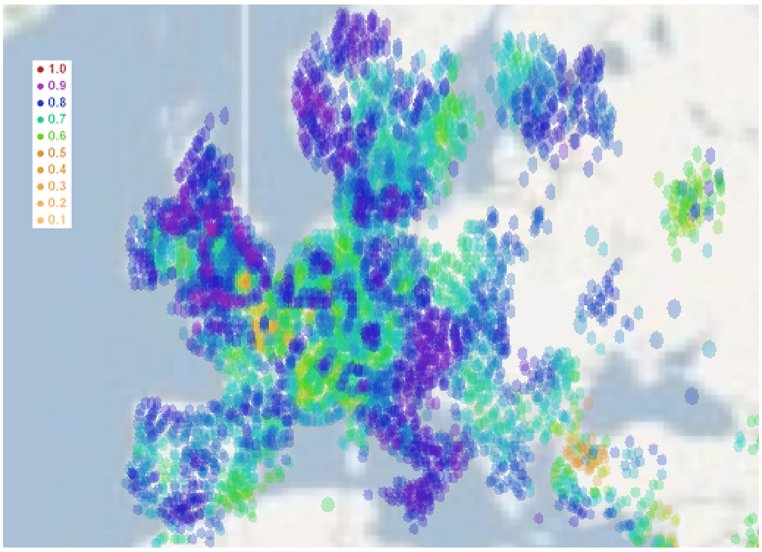

FIGURE 5.1: Western folk music interregional affinities
Europe with Askam-in-Furness, England.

- Figure 5.2 shows the interregional performance affinities between the Ashland, England region and Europe. In contrast to those interregional affinities of Askam with Europe, those of Ashland are significantly less pronounced. Scotland, Wales, and Sicily have the strongest affinities to Ashland.

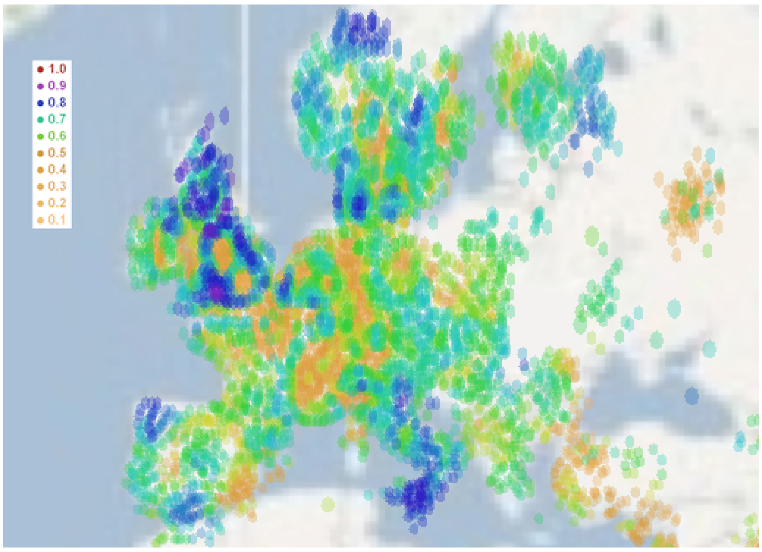

FIGURE 5.2: Western folk music interregional affinities
Europe with Ashland, Oregon, U.S.A.

- Figure 5.3 shows the interregional performance affinities between the Corigliano Calabro, Italy region and Europe. Again, we see a significant contrast of affinity strength as compared to Askam-in-Furness, but similar pockets of strength with Ashland, although slightly more pronounced.

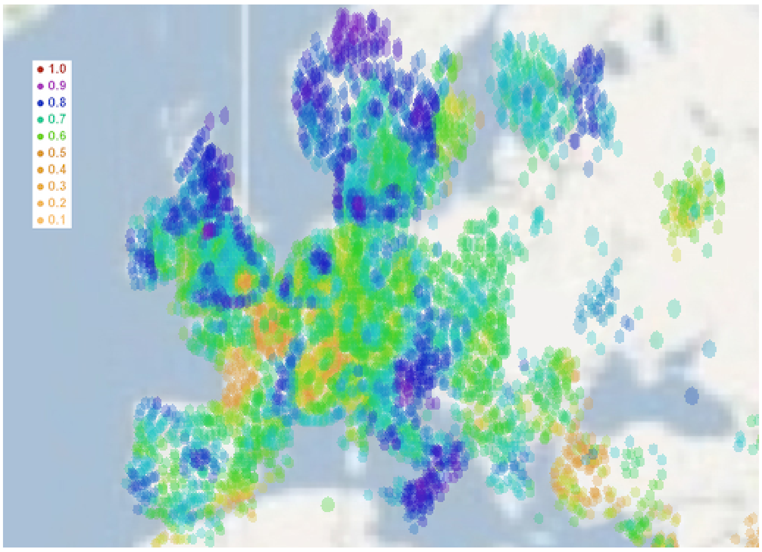

FIGURE 5.3: Western folk music interregional affinities
Europe with Corigliano Calabro, Italy.

There appear to be strong correlations in performance interpretation (again with respect to the use of time, and specifically note durations) between the Askam region and several regions throughout Europe, and substantially stronger than the affinities between Europe and the Corigliano or Ashland regions. Perhaps of interest, despite the overall higher concentration of affinities throughout Europe for Corigliano as compared with Ashland, some regions of Northern Portugal and Southern Spain have affinities approaching 0.80 with Ashland, but not so with Corigliano. In contrast, Corigliano performances have stronger affinities to southern Portugal than those of Ashland. Askam, however, has strong affinities with Northern Portugal, Southern Portugal, and Southern Spain. In other words, while most of the stronger affinities across Europe for the three regions are similar but with varying degrees of strength, there are regions in Europe that demonstrate stronger affinities with one region versus another.

5.3 case study: Southeast Asia

A Midsummer Night's Dream

Following the case study for Europe, we calculated interregional affinities for Western folk performances between the Askam-in-Furness, Ashland, and Corigliano regions and regions throughout Southeast Asia. The resulting interregional affinities are graphed below.

- Figure 5.4 shows the interregional performance affinities between the Askam-in-Furness, England region and those of Southeast Asia. Strong affinities exist in North/Central Thailand as well as parts of Malaysia, Indonesia, and the Philippines.

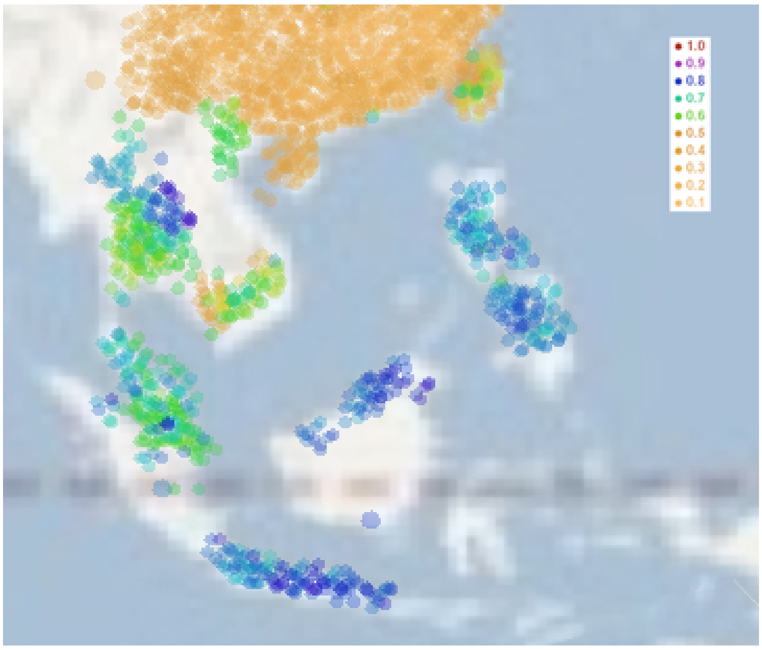

FIGURE 5.4: Western folk music interregional affinities Southeast Asia with Askam-in-Furness, England.

- Figure 5.5 shows the interregional performance affinities between Ashland, Oregon, U.S.A., and Southeast Asia. Strong affinities exist in the Southern Philippines, but unlike the Askam example, no obvious affinities exist in North/Central Thailand, Malaysia, or Indonesia.

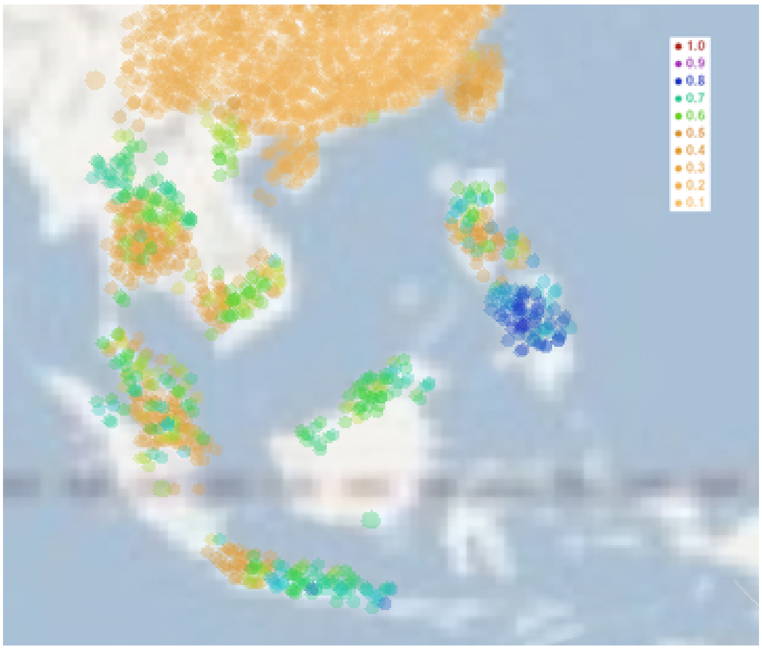

FIGURE 5.5: Western folk music interregional affinities
Southeast Asia with Ashland, Oregon, U.S.A.

- Figure 5.6 shows the interregional performance affinities between Corigliano Calabro, Italy and Southeast Asia. Moderate affinities are apparent in some sections of Indonesia and Malaysia.

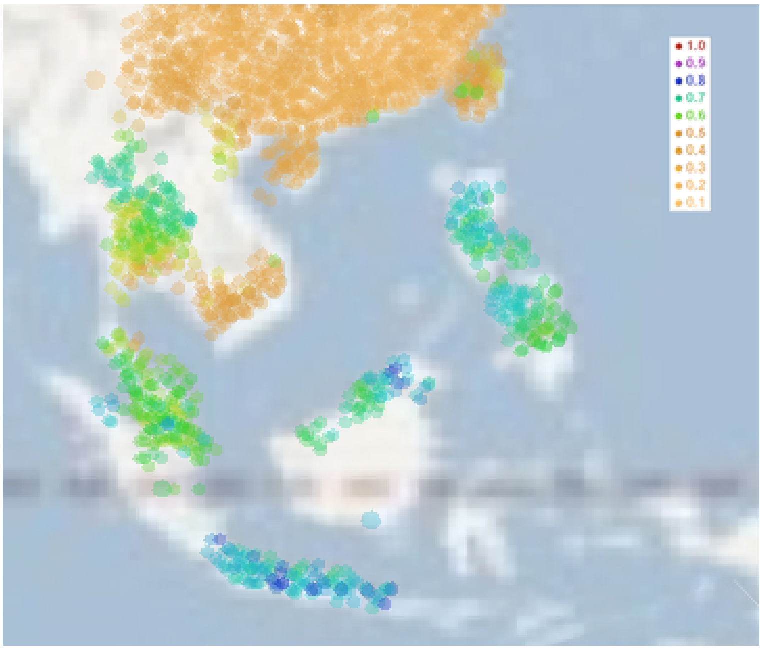

FIGURE 5.6: Western folk music interregional affinities
Southeast Asia with Corigliano Calabro, Italy.

Brisbane

Significant overseas born populations[67]

Country of birth	Population (2011)
United Kingdom	109,583
New Zealand	99,285
India	22,115
Mainland China	20,972
South Africa	19,587
Philippines	15,941
Vietnam	14,107
Germany	9,492
Malaysia	9,301
Fiji	8,568
United States	8,304
South Korea	8,210

Sydney

Significant overseas born populations[6]

Country of Birth	Population (2006)
United Kingdom	175,166
People's Republic of China	109,142
New Zealand	81,064
Vietnam	62,144
Lebanon	54,502
India	52,975
Philippines	52,087
Italy	44,563
Hong Kong	36,866
South Korea	32,124
Greece	32,021
South Africa	28,427
Fiji	26,928
Malaysia	21,211
Indonesia	20,562
Iraq	20,216
Germany	19,364
Sri Lanka	17,917
United States	16,340
Egypt	16,238
Croatia	15,501

Melbourne

Significant overseas born populations[1]

Place of Birth	Population (2006)
United Kingdom	156,457
Italy	73,801
Vietnam	57,926
China	54,726
New Zealand	52,453
Greece	52,279
India	50,686
Sri Lanka	30,594
Malaysia	29,174
Philippines	24,568
Germany	21,182
Malta	18,951
South Africa	17,317
Republic of Macedonia	17,287
Hong Kong	16,917
Poland	16,439
Croatia	15,367
Lebanon	14,645
Netherlands	14,581
Turkey	14,124
Bosnia and Herzegovina	13,546
Egypt	11,580
Total	774,600

FIGURE 5.7: Immigration sources for Brisbane, Sydney, and Melbourne (Wikipedia, "Australian Immigration Trends").

Like Europe, most strong regional affinities throughout Southeast Asia for performances of the Western folk music contained in this corpus pertain to the Askam-in-Furness, England region. With respect to the Philippines, performances from the Askam region showed strong affinities with those from both the North and the South. In contrast, performances from the Ashland region show strong affinities only with those regions from the Southern Philippines, while the Corigliano region has only weak affinities with the Philippines. For Southeast Asia the interregional affinities of Ashland and Corigliano are more or less inverted: there are no strong affinities between Corigliano and the Philippines for performances of the Western folk music, but moderately strong affinities with some sections of Indonesia and Malaysia. Ashland, on the contrary, has a strong affinity with the Philippines but only weak affinities with Indonesia and Malaysia.

5.4 case study: Christmas in Australia

Wie der Hirsch Schreit (Psalm 42)

Performances of Western Christmas music served as the basis for the interregional affinity analysis for the third case study. Using the statistical search method described above,

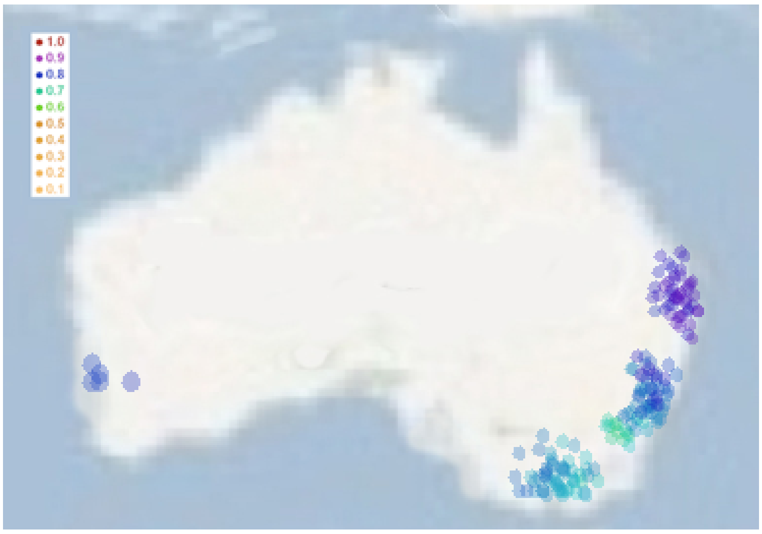

FIGURE 5.8: Western Christmas music interregional affinities Australia with Kemnay, Scotland.

the performances in the region encompassing Kemnay, Scotland (GPS 57.207108, -2.5) were found to have a high intraregional affinity, or 0.880. Performances from this region where then compared with those across the Australian continent. Australia was chosen because recent immigration trends are well documented. Figure 5.7 shows the principle sources of immigration for current inhabitants of the three largest cities in Australia, namely Brisbane, Sydney, and Melbourne (Wikipedia, "Australian Immigration Trends"). While all cities have recently experienced significant immigration, Brisbane's population is less diverse, with more immigration from the U.K. and New Zealand, whereas Sydney and Melbourne have had more substantial immigration from Asia, Europe, and the Middle East, as well as a higher percentage of recent immigrants among their respective populations.

The hypothesis for the interregional affinities between Kemnay, Scotland and Australia for performances of Christmas music: regions with less immigration and therefore of principle Anglo-Scot descent would share higher performance affinities for Western Christmas music with Scotland; regions with more immigration and diversity would demonstrate lower performance affinities with Scotland. Figure 5.8 graphs the relations of the interregional comparisons, which support the hypothesis. Regions surrounding Brisbane have the strongest performance affinities to those of Scotland, Melbourne the least, with Sydney somewhere in between the two extremes. There also appears to be a North/-South dichotomy of performance affinities with Scotland in Sydney, and an East/West dichotomy in Melbourne.

5.5 summary

Lieder ohne Worte

While the author of this research freely concedes the speculative nature of the interregional affinity analysis, the opportunity to explore global regional relationships through potential musical behaviors is intriguing. Of course several issues beset this analysis, including the question of statistical rigor. For example, while the regression analysis for calculating the *A-Pcc* (see Appendix A.3) demonstrated thresholds of significance, when exploring song genres that include multiple songs, the analysis of significance becomes quite complex. For this reason substantially more performances per sample (region) were included when calculating regional affinities across genres versus calculations for individual songs; this in turn would probably also lower the threshold for significance. Moreover, while a performance in Kemnay, Scotland certainly did emanate from that region — the performance was geo-tagged — little is known of the performer. The performer could be a missionary, an American oil worker, or even a visiting ethnomusicologist. Yet culture itself is not discriminatory. Certainly New York City's culture, to cite one, derives much from its many visitors and from the elevated degree of transience within its bounds. In other words, Sachs is correct; such is the evolution of culture.

Notwithstanding the limitations and potential flaws of this analysis, the fact that the interregional map of performance affinities in Australia reflects the immigration trends perhaps warrants further study. That there are differences in regional affinities in the Philippines for performances of Western Folk music, again, is not the expected outcome. The theory of identifying potential intrinsic cultural behaviors and norms through the analysis of geo-tagged amateur performance practices, while subject to doubt, holds promise. The overarching question confronting this theory, however, is whether such cultural practices are intrinsic or rather learned. This question, framed in terms of information viscosity, forms the basis of the next to last chapter.

Chapter 6

Viscosity

Csendes fehér tavat látok

6.1 Sachs

Riders on the Storm

> In pursuing our goal [the preservation of indigenous music], we are particularly rushed, as the venerable heritage of archaic cultures is threatened with imminent extermination. (Sachs)

The introduction of the iPhone, the iPad, and the Magic Piano application have likely hastened the very trend that Sachs feared. In this chapter, we confront the paradox (see Section 1.5) of studying music performances on mobile devices to identify intrinsic culture. Whereas, on the one hand, measuring intrinsic culture often requires access to technology (e.g. a phonograph in the case of Bartók and Kodály), on the other hand, that same technology quite likely accelerates the erosion of such cultures. Has this study identified intrinsic musical behaviors? Or instead, has this study simply measured the rate at which information, in this case predominantly Western music, has permeated the formerly protective regional membranes surrounding such intrinsic cultures?

6.2 background

Break On Through (To the Other Side)

In Chapter 3, we postulated that performance skill could be defined as knowledge, measured in the agogical enunciation of musical structure (see Section 3.5). Additionally, we proposed that such knowledge is acquired — that it is learned behavior. In Chapter 4, we found by partitioning the data by region, that some regions demonstrated knowledge of songs as measured through agogics, and that those regions are more likely to have homogeneous cultures. Moreover, those regions also showed significant intraregional affinities in performance. The level of knowledge or expertise, therefore, correlates directly with the level of performance interpretation similarities — the intraregional values. Hence the question: why or how does knowledge of a given song or genre exist within one region and not another? Does interpretative musical knowledge derive from intrinsic cultural norms? Or instead, are the knowledge and experience of a song or genre imported and assimilated?

Consider, once more, the results from Chapter 4 and the analysis of regional affinity by country (see Section 4.6). Table 4.9 is again referenced (see Table 6.1; for a listing of all intraregional performance affinities by genre and country, see Appendix D). Obvious gaps exist between both the regions *and* the genres. For example, performances in Saudi Arabia show a vast discrepancy of intraregional affinity in Western folk music as compared to video game music — a factor of eight. And intraregional affinities for performances of Western folk music in the Philippines are over six times greater than those in Saudi Arabia, and five times greater than those in China. Such vast gaps are surprising, and suggest that the data sample represents a culturally and socioeconomically diverse set of performers. More specifically, the gaps may also show that first, across diverse regions, there is also a high degree of cultural variability that impacts performance interpretation, and second, within a given region, for the intraregional affinities to vary across genre as they do, the cultural knowledge within the region also varies by song genre.

Region	Folk	Classical	Games	Xmas	Pop
Philippines	0.567	0.730	0.771	0.776	0.733
U.K.	0.474	0.625	0.697	0.769	0.725
U.S.A.	0.455	0.674	0.709	0.803	0.759
China	0.109	0.346	0.568	0.379	0.493
Saudi Arabia	0.084	0.568	0.700	0.482	0.633

Table 6.1: Intraregional affinities by genre and country.

6.3 viscosity theory

Yes, the river knows

One theory explaining the discrepancies of intraregional performance affinities focuses on the genre itself. Classical and even folk music might be expected to have more interpretative possibilities than, say, pop or video game music. The data support this theory; pop as a genre has the lowest degree of variance across regions (again as it relates to the treatment of time in performances); video game music and Western Christmas genres also show a lower degree of performance variance by region.

The most obvious theory accounting for the vast discrepancies between intraregional values by genre posits that some regions are more likely than others to know the genre. For example, intraregional affinities for Western Christmas music are highest in Europe and the U.S.A., and lowest in China and Saudi Arabia. Video game music, by contrast, shows similar intraregional affinities in Saudi Arabia and the U.S.A. or U.K. China is a somewhat challenging case for the model because of its size with respect to geography, population, and the resulting inevitable degree of diversity. Yet even so, the performers in China, by this basic theory of knowledge, are more familiar with video game music than Western Christmas music.

Is it possible that the knowledge associated with musical genres and particular songs — as opposed to cultural norms and intrinsic behaviors — can be measured through the analysis of 'skill' and intraregional performance affinities? And if so, while the knowledge may be derived from or influenced by cultural norms, if the knowledge flows into a region from abroad, is it possible that it may flow at varying rates depending on factors like socio-economic development, immigration, or even the presence and extent of globalizing technologies such as radio, the Internet, and the iPhone? In other words, might these factors, themselves related to culture, influence the viscosity of information flowing into such regions? Which isn't to say that knowledge 'flowing' at the same rate into two distinct cultures will have the same impact on both, for certainly the presence of cultural norms can influence not only the rate at which knowledge is disseminated but also how it is assimilated, whether accepted or rejected. What we are suggesting is that the rate of the flow of information, itself, may vary across region, and that varying levels of "information viscosity" will affect the degree to which performers assimilate certain songs or musical genres. For example, does the Great Chinese Firewall, a system designed by the Chinese government to limit the flow of certain Internet content into China, affect the conveyance of information, including cultural knowledge, from West to East in this region?

Of course, distinguishing between intrinsic and acquired cultural knowledge through the analysis of amateur music performances or otherwise seems potentially absurd. How could one know? And so, the analysis of skill/knowledge and intraregional affinities throughout this study may potentially measure either or both. Assume, for the present, that the analysis has failed to measure the former but instead has measured the latter, namely acquired or imported cultural knowledge. Viewing the data through *this* lens, then, while not revealing intrinsic cultural behaviors per se, does enable us to measure the viscosity of the flow of information, in this case predominantly Western music, from the West to regions across the world. Such a hypothesis more or less concedes Agawu's *us/them* dichotomy.

> One of the most persistent and at the same time controversial dichotomies used by ethnomusicologists is the us/them construct. The construct aims at recognizing the differences between researcher and researched, between 'those who seek knowledge about other musical traditions and those that impart the knowledge', and between colonizer and colonized. (Agawu)

6.4 viscosity examples

Take it as it comes

Accepting this alternative theory, we take a second look at the intraregional performance affinity graphs for various musical genres in Asia, this time as indicative of viscosity of information. Such a theory presumes the intraregional affinities indicate the presence of knowledge of a song genre in a given region.

- Figure 6.1 shows the intraregional performance affinities for Western Folk music in Asia. Little cultural knowledge of Western Folk music is indicated — the genre is extremely viscous in Asia.

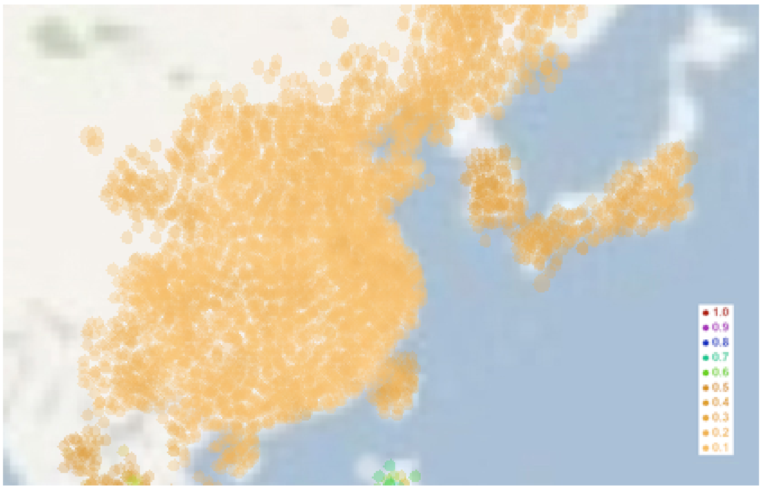

FIGURE 6.1: Intraregional performance affinities for Western folk music in Asia.

- Figure 6.2 shows the intraregional performance affinities for Western Classical music in Asia. Pockets of knowledge are present in China and Taiwan. In China the pockets seem to correspond with the primary urban areas in or near Beijing. Korea and Japan, unlike China and Taiwan, show significant knowledge of the genre, and thereby a lower level of viscosity for Western Classical music as compared with Western folk music.

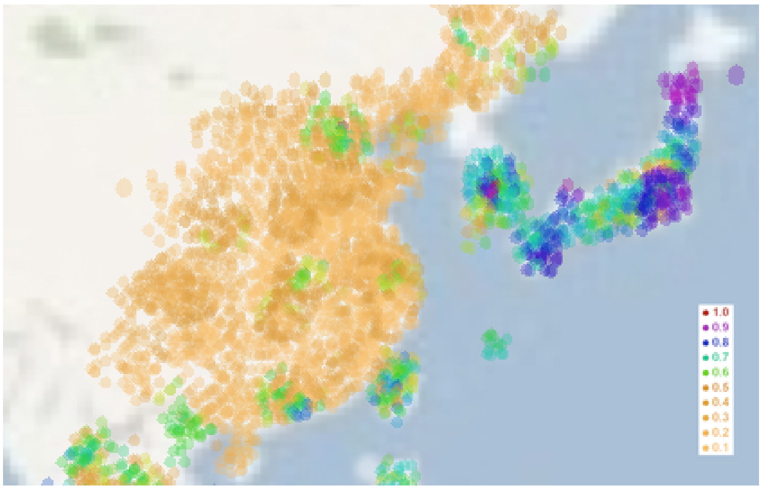

FIGURE 6.2: Intraregional performance affinities for Western Classical music in Asia.

- Figure 6.3 shows the intraregional performance affinities for video game music in Asia. Pockets of knowledge, once again, are present in China. Taiwan, Korea, and Japan indicate a high level of knowledge of this genre.

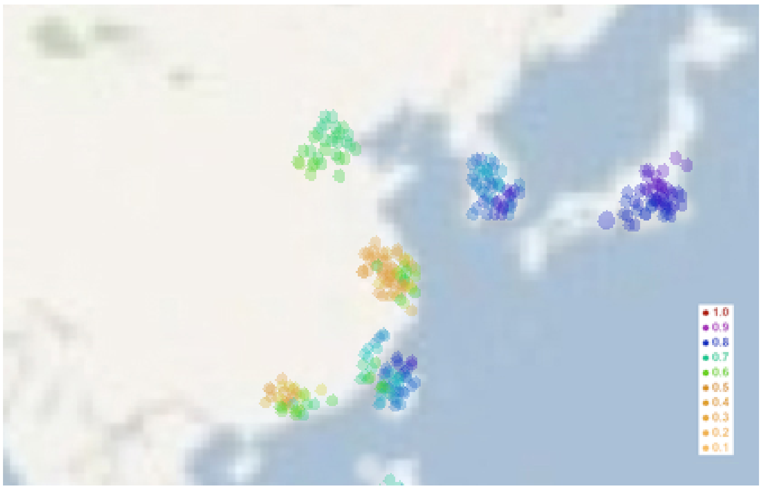

FIGURE 6.3: Intraregional performance affinities for video game music in Asia.

- Figure 6.4 shows the intraregional performance affinities for Western pop music in Asia. The graph suggests that Western pop is the least viscous form of information permeating China, as compared with other musical genres.

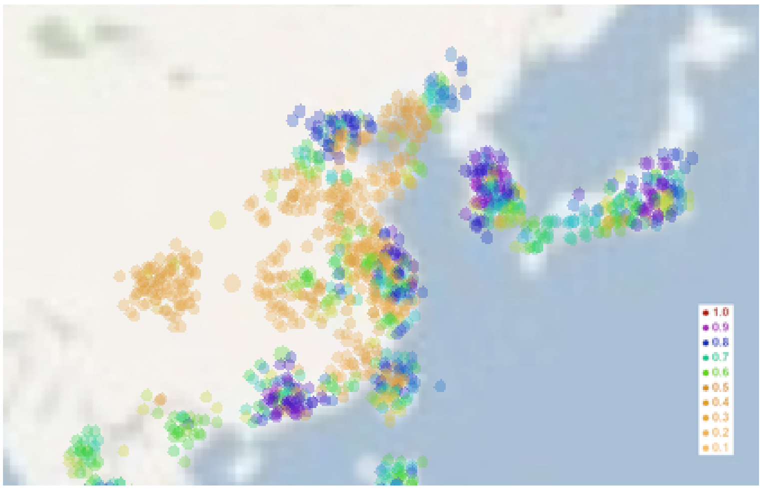

FIGURE 6.4: Intraregional performance affinities for Western pop music in Asia.

- Figure 6.5 shows the intraregional performance affinities for Western Christmas music in Asia. Like the Western folk music genre, Western Christmas music is quite viscous in China. But unlike Western folk music, it is less viscous in Japan and Taiwan, and less so again in Southern Korea.

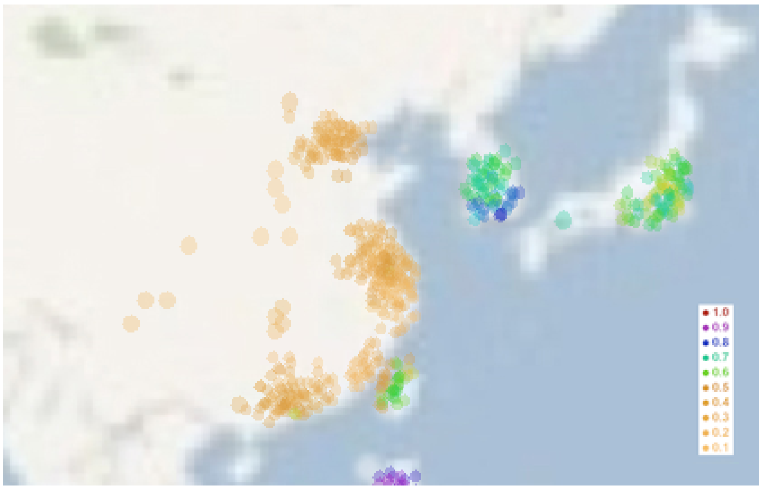

FIGURE 6.5: Intraregional performance affinities for Western Christmas music in Asia.

Once again, the author freely concedes the speculative nature of our analysis of intraregional performance affinity. But there does seem to be solid support for this alternative interpretation of the data. While on the one hand, a set of cultural norms in China may influence the musical interpretation of Western folk or Classical music, on the other hand, it's certainly seems possible that the data reflects the extent to which such musical knowledge has flowed into the region from abroad. Perhaps it reflects both cultural norms and viscosity. But whatever the case, it does not seem completely implausible that in China Western Christmas music is more viscous than Western pop music. Even putting aside all theories as to how to interpret the results, the data and analysis are sufficiently interesting to justify further study.

Chapter 7

The Seventh Door

Lásd a réegi asszonyokat

Fremd bin ich eingezogen,
Fremd zieh' ich wieder aus.
Der Mai war mir gewogen
Mit manchem Blumenstrauß.

Wilhelm Müller

7.1 Bartók

Die Nebensonnen

Perhaps this research has successfully identified intrinsic musical behaviors. Or perhaps it has done little more than confirmed Sach's fears. Yet, once again following Bartók, a third alternative is possible.

> Contact with foreign material not only results in an exchange of melodies, but — and this is still more important — it gives an impulse to the development of new styles. At the same time, the more or less ancient styles are generally well preserved, too, which should further enhance the richness of the music. The trend toward transformation of foreign melodies prevents the internationalization of the music of these peoples. The material of each, however heterogeneous in origin, receives its marked individuality. The situation of folk music in Eastern Europe may be summed up thus: as a result of uninterrupted reciprocal influence upon the folk music of these peoples

> there are an immense variety and a wealth of melodies and melodic types. The 'racial impurity' finally attained is definitely beneficial. (Bartók)

> It is obvious that if there remains any hope for the survival of folk music in the near or distant future (a rather doubtful outcome considering the rapid intrusion of higher civilization into the more remote parts of the world), an artificial erection of Chinese walls to separate peoples from each other bodes no good for its development. A complete separation from foreign influences means stagnation: well assimilated foreign impulses offer possibilities of enrichment. (Bartók)

7.2 conclusion

Gute Nacht

Although Aarden and Huron's theories hold promise, and the availability of a new and significant corpus of geo-tagged amateur performances provide new data for research, the search for universal behaviors remains elusive.

Whether intrinsic behaviors have been identified, or whether, instead, Sachs' fears and Agawu's dichotomy have been confirmed remains an open question. It is possible that this study achieved little beyond the viscosity of information flow of predominantly Western music across the globe. However, it is believed that the work, if indecisive in terms of providing an unambiguous theory regarding universality, challenges long held assumptions and opens new directions for studying human behavior in music performance. It is hoped that the data, the methods, and the theories postulated will enhance future research, and that the search for the metaphorical Transylvanian village of Bartók will continue.

Appendix A

Statistical Correlation Model

A.1 Pearson

The *Pearson correlation coefficient*, or *Pcc*, allows the statistical comparison of two sequential sets of numbers to ascertain the degree of correlation between them (Rodgers and Nicewander). This statistical method has proven effective in other research on musical timings. For example Sapp used the *Pearson correlation coefficient* to identify similarities (including fraudulently produced similarities) between expert performances of Chopin's *Mazurkas* (Sapp, "Comparative Analysis of Multiple Musical Performances"; "Hybrid Numeric/Rank Similarity Metrics for Musical Performance Analysis"). In the case of this research, the *Pearson correlation coefficient* was used to compare the note durations of two performances.

The application of the *Pearson* produces a coefficient ranging from

$$-1.0 \leq 0 \leq 1.0 \tag{A.1}$$

where

-1.0 indicates negative correlation;

0.0 indicates no correlation; and

1.0 indicates identity.

The following implementation of the formula was used in the study

$$r = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{[\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2]^{\frac{1}{2}}} \tag{A.2}$$

In the numerator, the raw scores are centered by subtracting out the mean of each variable, and the sum of cross-products of the centered variables is accumulated. The denominator adjusts the scales of the variables to have equal units. Thus Equation (A.2) describes r as the centered and standardized sum of cross-product of two variables. (Rodgers and Nicewander)

The following *java* code was used to implement the formula:

```
public double PearsonCorrelation(Performance Y)
{
        double mean = 0;
        double x_stddev = 0;
        double y_stddev = 0;
        for (int i = 0; i < numNotes; i++) {
                double x = noteDur.get(i) - noteAverageDuration;
                double y = Y.noteDur.get(i) - Y.noteAverageDuration;
                mean += x * y;
                x_stddev += x * x;
                y_stddev += y * y;
        }

        double stddev = Math.sqrt(x_stddev) * Math.sqrt(y_stddev);
        double r = mean / stddev;

        return r
}
```

Listing A.1: Pearson Correlation Java Code

A.2 Average Pearson

The average *Pearson correlation coefficient*, or *A-Pcc*, is used throughout the study to assess the similarity between performances of one set and itself or of one and other sets of performances. To calculate the average *Pearson correlation coefficient* within a set of performances, each performance within the set is compared to every other performance in the set via *Pearson*. The total coefficient is then divided by the number of comparisons to produce the average, which, like the coefficient itself, would range from $-1 \leq 0 \leq 1$.

$$A{-}Pcc = \frac{\sum_0^n r}{n} \tag{A.3}$$

The following *java* code was used to calculate the *A-Pcc* for two sets of performances, A and B:

```
public double CalcAvePearson(MPSongList A, MPSongList B)
{
        long count = 0;
        double average = 0;
        for (MPPerformance a : A.songs) {
                for (MPPerformance b : B.songs) {
                        if (!a.equals(b)) {
                                average += a.PearsonCorrelation(b);
                                count++;
                        }
                }
        }
        return average / (double)count;
}
```

Listing A.2: Average Pearson Correlation Java Code

To calculate the *A-Pcc* of a set, A, to itself, the following *java* code was used:

```
double r = CalcAvePearson(A, A);
```

Listing A.3: Intraregional Affinity Java Code

This calculation yields an *intraregional* affinity of performances within region A. To calculate the *interregional* affinity between two different regions, namely region A and region B, the same code is used with the second parameter changed:

```
double r = CalcAvePearson(A, B);
```

Listing A.4: Interregional Affinity Java Code

A.3 Pearson significance and sample size

The number of performances within a set that are compared to one another in order to calculate similarities via the *Pearson correlation coefficient* will obviously impact the coefficient. The following graph (see Figure A.1) shows the relationship between the size of a set and the threshold of significance (Wikipedia, "Correlation Significance"). The

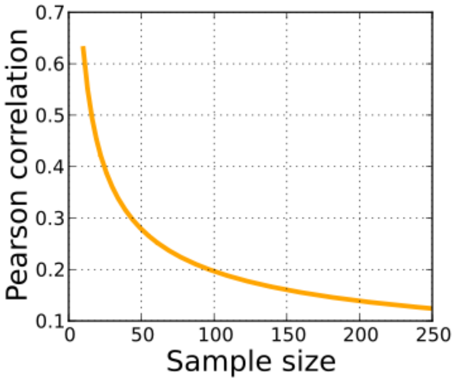

FIGURE A.1: Pearson significance by sample size.

graph proposes that as the sample size contracts, higher coefficients are required in order to deem the correlation significant. The graph also implies that as the set (in this case performances) grows, an inevitable gravitation towards a mean reduces the likelihood of strong correlations within the set, and hence a much lower threshold of significance.

Some basic regression analysis was performed with the performance data in order to ascertain optimal sizes of sets of performances which might be compared, both to themselves to learn of similarities within the sets (i.e. internal affinity) as well as to other sets in order to identify potential similarities across sets (i.e. external affinity). To run the regression tests, several hundred thousand performances across multiple song genres were divided into approximately 3,900 distant geographical regions of approximately 120 square kilometers each. The *Average Pearson correlation coefficient* (see above) was computed for all regional sets. The correlation between the size of the sets and the average internal affinity of the sets was then calculated. Figure A.2 graphs the *A-Pcc* by minimum region size — all regions with performances greater than or equal to this minimum number of performances. The graphs suggests inflection points at 20 and 40.

Based on the significance standards and this regression analysis, this study used set sizes of between 24 and 26 performances per region for most of the inter and intra regional analysis, yielding the so called external and internal affinities as measured by the *A-Pcc*. During much of the regional analysis, when a given region yielded fewer than the minimum target of 24 performances that region was either excluded or iteratively enlarged (geographically) to produce the minimum target number of performances. Such iterative expansion of the area of the region was only accepted to various thresholds; for example, a regional radius of 70 km was only allowed to grow to a maximum radius of 300 km. Such expanded regions would generally be extremely remote and isolated. The software also allowed a region with greater than the target number of performances to be recursively subdivided, yielding smaller regions with fewer performances, or within

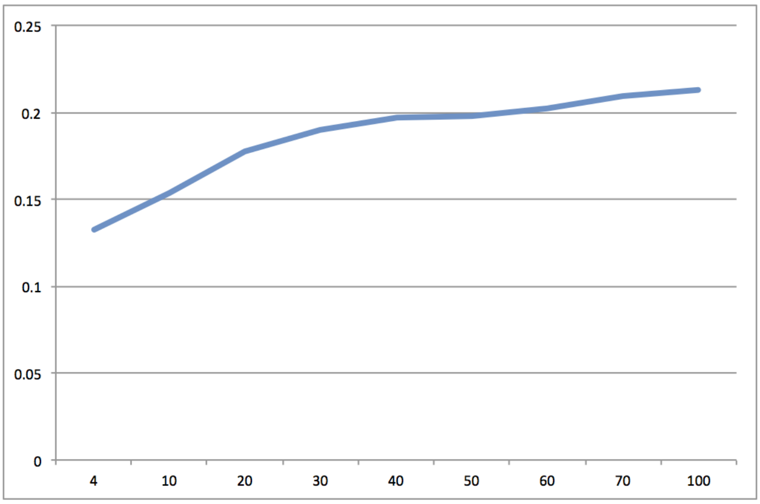

FIGURE A.2: Minimum sample size and intraregion A-Pcc.

the target maximum. Such regions would represent densely populated urban areas; for example, mid-town Manhattan might be split at 5th, and then again at 3rd and 8th, etc. In the case where the software evaluated song genres instead of individual songs, minimum performance thresholds per region were correspondingly increased. All of these thresholds and targets are parameters in the analysis software, and can be freely changed to any arbitrary value. To validate the significance of the results, the minimum and maximum targets were tested with multiple values. These tests confirmed the parameters established for this study (e.g., reducing the minimum produced an unacceptable standard of deviation; increasing the maximum did not increase precision without migrating a set towards a mean, thereby reducing the intrinsic affinities endemic to the set).

Appendix B

Magic Piano Data Collection

B.1 background

Users of the "Magic Piano" mobile application have created hundreds of millions of performances of piano music. As discussed in Chapter 2, the research examined performances of over 300 songs across several genres (see Section C). The vast majority of the songs were of Western origin. Figure B.1 shows the location of the performances in the corpus.

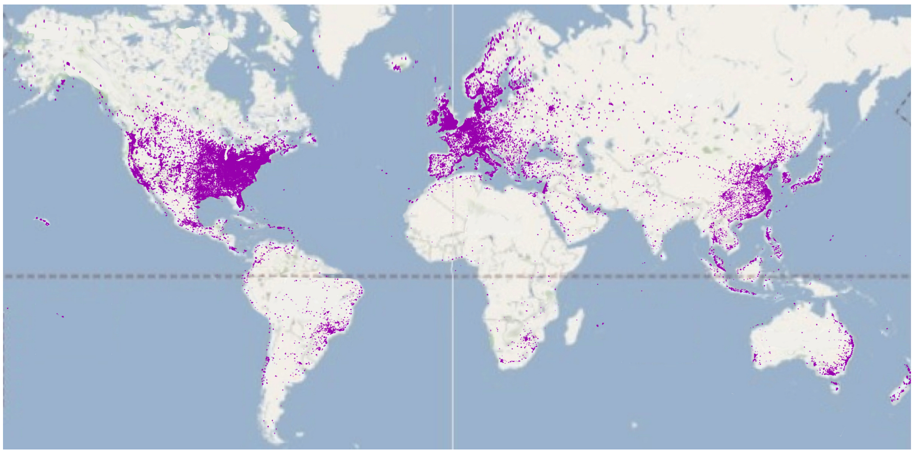

FIGURE B.1: Location of each performance in the data set.

B.2 approach

Given the findings of Schaffrath, the following principles have been adopted (Schaffrath):

- Integrity of all musical data has been preserved. The *MIDI* representation of each song is stored and presented exactly as produced by the application.
- GPS coordinates and time stamp are given for each performance.
- All data is stored using non-proprietary, open standard representations:

  All musical data is presented in the MIDI 1.0 format;

  All metadata, including GPS coordinates, is presented in a JSON format (Javascript Open Notation), a light-weight text-based, human-readable representation designed for interchange.
- A catalog of the data is offered in a hierarchical JSON structure, allowing consistent and clear identification of song titles and performance parameters.
- The entire corpus of data is available read-only via standard HTTP and FTP protocols.
- Sample *java* code to read and manipulate the data has been published alongside the data.
- Roughly 800,000 performances were published.

B.3 data

The data is published in the Digital Archive of (Mobile) Musical Performances (or *DAMP*) on Stanford's CCRMA site.

```
https://ccrma.stanford.edu/damp/magicpiano/
```

Listing B.1: Location of corpus

The data is organized by directory of genre and song. For example, the Western classical performances are here:

```
https://ccrma.stanford.edu/damp/magicpiano/perfs/classical/
```

Listing B.2: Location of classical music performances

For each performance of each song in each genre, there is the native *MIDI* and a corresponding *JSON*. The *JSON* contains meta data for the performance, including the date, time, and location of the performance. The location is given both as GPS coordinates and as reverse-geocoded textual location data, including, for example, the name of the city, country, etc.

The *JSON* meta data is stored as follows:


```
{
    "GPS": "40.9039,14.0579",
    "Location": {
        "address": "Via Ripuaria, 80014 Varcaturo Province of Naples, Italy",
        "administrativeAreaName": "Campania",
        "commonLocalityName": "Naples",
        "countryName": "Italy",
        "countryNameCode": "IT",
        "is_city": true,
        "localityName": "Varcaturo",
        "postalCodeNumber": "80014",
        "thoroughfareName": " Via Ripuaria"
    },
    "date": "Oct 14, 2012",
    "user-id": 25722966
}
```


Listing B.3: Example JSON performance meta data

The GPS data is given in lat/long format. If the reverse geo-code failed for the GPS data (we attempted to reverse geo-code all coordinates with Google's APIs, but a few queries failed to return a valid location), various parameters of the "Location" field could be empty. The user-id is unique across all performances, and hence could be used to identify repeat performances from the same user for various songs, of course noting that the location may vary across such performances (e.g. they played a song at their home and then at their work location). The is_city field is a function of the parameters currently set for the research, and may or may not reflect other definitions of an urban area. As noted above, we used a radius of 0.35 degrees (or roughly 40 kilometers) for the 760 largest urban areas of the world to set this value (see the *RadiusMatch* function in the *MPCityItem* class in the *JAVA* source code).

The following shows the location of performances of Chopin Prelude Op 28 No. 6 and the corresponding *JSON* files for this song:

```
https://.../classical/chopin_prelude_op28no6/D19DA658-A784-CCB832AFC9EB.midi
```

```
https://.../classical/chopin_prelude_op28no6/D19DA658-A784-CCB832AFC9EB.json
https://.../classical/chopin_prelude_op28no6/0A704CDE-9D32-2EB872C4CF4B.midi
https://.../classical/chopin_prelude_op28no6/0A704CDE-9D32-2EB872C4CF4B.json

etc.
```

Listing B.4: Location of Chopin Prelude Op 28 No. 6 performances

Where possible, there is the *MIDI* file of the reference performance which the *Magic Piano* application uses to create the firefly cues. Such data are useful to facilitate analysis; the reference *MIDI* file, for example, was often used to compute *slope* for the *AAT* of a performance (or set of performances). There is also a *JSON* file that includes parameters used by the *Magic Piano* application, including, for example, the initial rate at which the fireflies descend. These files are stored in the following directory:

```
https://ccrma.stanford.edu/damp/magicpiano/orgperfs/
https://ccrma.stanford.edu/damp/magicpiano/orgperfs/sakura_sakura.midi
https://ccrma.stanford.edu/damp/magicpiano/orgperfs/sakura_sakura.json
https://ccrma.stanford.edu/damp/magicpiano/orgperfs/amazing_grace_bundled.midi
https://ccrma.stanford.edu/damp/magicpiano/orgperfs/amazing_grace_bundled.json

etc.
```

Listing B.5: Location and example baseline performance and configuration file

B.4 samplecode

All of the *JAVA* code used in the analysis for this research is published here:

```
https://ccrma.stanford.edu/damp/magicpiano/source/
```

Listing B.6: Sample source code

The code has limited documentation. The curious reader is therefore advised to heed the admonishments of Darth Vader: “Use the force, read the source.” The code includes a sample project in Eclipse that was used to build and run the project, as well as associated configuration and support files.

Appendix C

Song listings

adagio glass harmonica	adagio winter	air g
album leaf	allegro primavera	appalacian spring
aria bach	aria dmin	aria dmin scarlatti
ave maria biebl	ave maria gounod	bach invention 4
bach toccata dm	bagatelles opus 126 andante con moto	bartok for children 2
bartok for children 5	bourree en rondeau	bourreeii amaj
brahms lullaby	bydlo	childrens corner
chopin etude no 9	chopin etude op10no3	chopin prelude no20
chopin prelude no 15	chopin prelude no 4	chopin prelude op28no12
chopin prelude op28no13	chopin prelude op28no5	chopin prelude op28no6
chopin prelude op28no7	chopin prelude op28no9	claire de lune
confutatis	contrapunctus ix	eine kleine nachtmusik
etude 25 no 1 harp study	etude in e min	flight of the bumblebee
fremdenlandern	fugue dm	fur elise bundled
gallop marquis	gavotte II DMaj	german dance BbMaj
german dance Fmaj	gigue in g	gnomus
goldberg variation no 5	golliwogs cakewalk	gymnopedie no 1
gypsy dance	hall of the mountain king	halleluja messiah
humming song	ich grolle nicht	invention no 1
invention no 13	jesu joy of mans desiring	klavierstuck FMaj
kleiner trauermarsch c	kyrie	la fille aux cheveux de lin
lacrymosa	lamour oiseau	largo handel
maple leaf rag	march in eb major	march prokofiev
menuett c	menuett d	minuet boccherini
minuet eb maj	minuet in g	minute waltz
moonlight ii allegretto	moonlight iii presto agitato	moonlight sonata
morning mood from peer gynt	musette DMajor	nutcracker march easy
nutcracker march hard	ode to joy	old castle
pachelbels canon	parsley and celery	partita1 allemande
partita1 courante	partita1 gigue	partita1 minuet1
partita1 minuet2	partita1 prelude	partita1 sarabande
piano quintet eb major	piano sonata pathetique	pictures at an exhibition
pictures at an exhibition easy	pomp and circumstance 2	prelude c minor
prelude no1 in cmaj	rach prelude csharp min	rach prelude csharp minor
ravel pavane	requiem	reverie debussy
rhapsody paganini	rondo alla turca	scriabin prelude gmin
sheep may safely graze	sonata 8 rondo	sonata gmaj
sonatina no 1	song of the lark	songs without words 19 1
the bell tolls	the entertainer	the little shepherd
the tango endless	to a wild rose	traumerei
tuba mirum	venetian gondolier	waltz bmin
waltz emaj	wedding march	wedding march mendelssohn
wohin		

Table C.1: Listing of classical songs.

100 years	SOS	a thousand miles
addams family theme	airplanes bob	all by myself
all the right moves	all the small things	all the things you are
allentown	always a woman	always be my baby
american girl	and i am telling you im not going	another saturday night
apologize	as long as you love me	baby bieber
baby one more time	back on the chain gang	bad romance
beauty and the beast	before he cheats	billionaire
birdhouse in your soul	bootylicious	borderline
born this way	bottle it up	boyfriend
breakeven	brick	bring me to life
bubbly	burnin up	california dreamin
california sun	call me maybe	call me when youre sober
cant let go mariah carey	cheers	clocks
close every door	come sail away	cooler than me
cough syrup	count on me	country honk
daddy sang bass	dance dance	dancing in the street
daydream believer	devil went down to georgia	disturbia
do it again	dont cry for me argentina	dont fear the reaper
dont get me wrong	dont know why	dreamlover
drive	drive incubus	drops of jupiter
dynamite	endless love	every rose has its thorn
everybody rock your body	everybody talks	fairytale of new york
fantasy mariah carey	fireflies	first date
first kiss	fix you	flightless bird american mouth
forget you	free bird	friend like me
georgia on my mind	get the party started	gimme shelter
gimme three steps	glad you came	good time
gravity	great balls of fire	grenade
hard to handle	heart alone	heart of glass
here comes your man	hero	hey soul sister
hit me with your best shot	home	home phillip phillips
house of the rising sun	how to save a life	how you remind me
i believe i can fly	i dont want to be	i dont want to cry
i just died in your arms tonight	i look to you	i wanna dance with somebody
i wanna go	i want it that way	i want to know what love is
i want you back	i will survive	i wont give up
id do anything for love	if i were a boy	ifihadamilliondollars
ill be there	im yours	in a sentimental mood
in love with a girl	interstate love song	istanbul
it will rain	its been awhile	just dance
just the way you are bm	just what i needed	king of anything
kiss the girl	la bamba revised	lady marmalade
landslide	larger than life	lean on me
let her cry	let me love you	losing my religion
love game	love shack	love song
lovebug	lucky	lucky bs
mad world	man in the mirror	marry you
misery	more than words	moves like jagger
movin out	moving in stereo	mr cellophane
my heroes have always been cowboys	my life	my life would suck without you correct
my own worst enemy	my way	need you now
no air	not over you	nothin on you
nothing	on the floor	one angry dwarf
one less lonely girl	one of us	one thing
one week	only the good die young	oops i did it again
paparazzi	particle man	piano man
pinch me	please dont go	poker face
pound the alarm	praan	pty mj
ramblin man	respect	rhythm of love
rock and roll all nite	rocky mountain way	russian roulette
safety dance	santeria	save the last dance for me
scarborough fair	shape of my heart	she will be loved
shes so high	since u been gone	so far away
somebody to love bieber	somewhere in brooklyn	son of a preacher man
stand by me	stronger	sugar were goin down
super bass	superman	surfin safari
sweet home alabama	take on me	take the heartland
thats the way it is	the entertainer bj	the lazy song
the way i am	this aint a scene its an arms race	total eclipse of the heart
u drive me crazy	u smile	unpretty
viva la vida	walking in memphis	what the hell
whats my age again	where is my mind	wild rover
wild thing	wind beneath my wings	wont go home without you
wrong way	yellow	you and i
you are not alone	you lie	you make loving fun
you raise me up		

Table C.2: Listing of pop songs.

away in a manger	deck the halls
God rest you merry gentleman	hark the herald
jingle bells	joy to the world
o christmas tree	o come all ye faithful
o holy night	o little town of bethlehem
the first noel	we three kings

Table C.3: Listing of Christmas songs.

a whole new world	all i ask of you	another suitcase in the hall
any dream will do	beauty and the beast	cabaret
can you feel the love tonight	castleonacloud	circle of life
close every door	colors of the wind	dont cry for me argentina
friend like me	hakuna matata	i dont know how to love him
i dreamed a dream	i got rhythm	i just cant wait to be king
kiss the girl	luck be a lady	make believe
mr cellophane	ol man river	once upon a dream
over the rainbow	part of you world	phantom of the opera
pinkpanther	reflection	seasons of love
shall we dance	someone like you	spoonful of sugar
superstar	take me or leave me	the bare necessities
under the sea	you must love me	zip a dee doo dah

Table C.4: Listing of show tunes.

castlevania2 day	castlevania3	chronotrigger frog
dkc jungle	dr mario	dragon warrior main
ff battle medley	ff boss medley	final fantasy prelude
king k rool	mario brothers	mario underwater
sonic	super mario bros 2	tetrisA
tetrisB	wily megaman2	zeldatheme

Table C.5: Listing of video game songs.

amazing grace	america the beautiful	appalacian spring
auld lang syne	battle hymn of the republic	coming round the mountain
frere jacques	go tell it on the mountain	God save the queen
good morning to all	joy to the world	man of constant sorrow
my old kentucky home	o canada	sakura sakura
shenandoah	star spangled banner	telephone polka
the eyes of texas	the farmer in the dell	the irish washerwoman
turkey in the straw	twinkle twinkle bundled	when irish eyes are smiling
when the saints go marching in	whiskey in the jar	wild rover

Table C.6: Listing of folk songs.

Appendix D

Intraregional Affinity Tables

Approximately 800,000 performances of mostly Western folk, pop, classical, holiday, show, and video game songs (see Appendix C) were partitioned and analyzed by country. In addition to the intraregional affinities, the interregional affinities were also calculated, namely the level of similarity between each performance in a region (in this case country) and every other region/country. Only countries with a significant number of performances were considered (see Appendix A.3). The results, partitioned by genre, were then sorted and ranked by country.

Country	Average Intraregional affinity	Average Interregional affinity
Germany	0.890	0.861
South Korea	0.884	0.854
Malaysia	0.880	0.860
Belgium	0.870	0.841
Singapore	0.869	0.851
Philippines	0.868	0.844
Thailand	0.863	0.842
Russia	0.863	0.853
Canada	0.861	0.848
South Africa	0.861	0.850
New Zealand	0.860	0.848
Japan	0.858	0.842
France	0.855	0.840
Spain	0.853	0.839
Brazil	0.852	0.846
Australia	0.851	0.843
China	0.849	0.839
Taiwan	0.848	0.830
Italy	0.846	0.838
Indonesia	0.840	0.828
Mexico	0.829	0.829
The Netherlands	0.826	0.836
Denmark	0.825	0.826
Sweden	0.824	0.825
UK	0.824	0.821
USA	0.816	0.818
Norway	0.811	0.826

Table D.1: Intraregional affinity for show songs.

Country	Average Intraregional affinity	Average Interregional affinity
Indonesia	0.809	0.733
Slovakia	0.801	0.676
Iceland	0.770	0.646
Belarus	0.756	0.666
Iran	0.753	0.663
Japan	0.753	0.694
South Korea	0.747	0.695
Singapore	0.738	0.701
Croatia	0.735	0.616
Georgia	0.733	0.644
Philippines	0.730	0.690
Chile	0.728	0.697
Czech Republic	0.723	0.644
Malta	0.721	0.641
Panama	0.709	0.584
Turkey	0.708	0.696
Spain	0.708	0.698
Norway	0.704	0.672
Germany	0.703	0.681
Azerbaijan	0.700	0.619
Malaysia	0.699	0.686
Bulgaria	0.697	0.614
Hungary	0.696	0.608
Switzerland	0.695	0.678
Belgium	0.693	0.682
New Zealand	0.692	0.685
Sweden	0.687	0.681
The Netherlands	0.687	0.666
Morocco	0.686	0.550
Guam	0.685	0.588
France	0.685	0.665
Vietnam	0.681	0.664
Russia	0.680	0.662
Macau	0.680	0.567
Italy	0.676	0.669
United States	0.674	0.611
Brazil	0.668	0.655
Taiwan	0.667	0.669
Peru	0.667	0.588
Denmark	0.665	0.671
Portugal	0.661	0.621
Honduras	0.658	0.545
Colombia	0.657	0.660
Dominican Republic	0.653	0.566
Australia	0.652	0.647
Ecuador	0.650	0.592
Canada	0.648	0.640
Ukraine	0.645	0.614
Guatemala	0.643	0.603
Venezuela	0.643	0.605
Mongolia	0.643	0.567
Greece	0.639	0.592
South Africa	0.628	0.653
UK	0.625	0.609
Thailand	0.616	0.641
Iraq	0.615	0.572
Finland	0.610	0.604
Kazakhstan	0.602	0.577
Jordan	0.598	0.555
Brunei Darussalam	0.594	0.563
Austria	0.592	0.572
Egypt	0.581	0.575
Mexico	0.575	0.586
Saudi Arabia	0.568	0.599
Romania	0.564	0.579
Hong Kong	0.562	0.573
Poland	0.558	0.557
Argentina	0.556	0.578
United Arab Emirates	0.550	0.574
Bahrain	0.513	0.502
Puerto Rico	0.501	0.543
Ireland	0.488	0.535
Costa Rica	0.476	0.531
Israel	0.448	0.534
Lebanon	0.401	0.492
Kuwait	0.396	0.471
Qatar	0.378	0.473
China	0.346	0.460
India	0.331	0.439

Table D.2: Intraregional affinity for Western classical songs.

Country	Average Intraregional affinity	Average Interregional affinity
United States	0.803	0.755
Malaysia	0.798	0.739
Norway	0.794	0.736
Indonesia	0.778	0.725
Philippines	0.776	0.728
Australia	0.776	0.754
South Africa	0.775	0.723
United Kingdom	0.769	0.732
Canada	0.769	0.743
Sweden	0.768	0.722
Mexico	0.741	0.736
Germany	0.740	0.714
Italy	0.738	0.737
The Netherlands	0.730	0.709
New Zealand	0.708	0.694
Denmark	0.708	0.699
Brazil	0.702	0.693
Spain	0.699	0.693
Colombia	0.676	0.675
Thailand	0.668	0.672
Russia	0.667	0.679
Japan	0.666	0.676
South Korea	0.654	0.665
France	0.628	0.654
Saudi Arabia	0.482	0.570
China	0.379	0.546

Table D.3: Intraregional affinity for Christmas songs.

Country	Average Intraregional affinity	Average Interregional affinity
Japan	0.823	0.768
Thailand	0.813	0.769
Taiwan	0.800	0.763
Belgium	0.787	0.757
Spain	0.786	0.757
South Korea	0.780	0.752
France	0.774	0.749
Germany	0.773	0.749
Philippines	0.771	0.741
Denmark	0.770	0.740
Malaysia	0.769	0.750
Singapore	0.763	0.743
Indonesia	0.758	0.737
Chile	0.749	0.714
Italy	0.745	0.741
Russia	0.745	0.731
Sweden	0.736	0.731
Australia	0.732	0.728
Norway	0.728	0.730
Canada	0.725	0.713
The Netherlands	0.723	0.729
USA	0.709	0.705
Colombia	0.705	0.711
Saudi Arabia	0.700	0.709
UK	0.697	0.701
Turkey	0.677	0.698
Brazil	0.665	0.692
Mexico	0.644	0.683
China	0.568	0.650

Table D.4: Intraregional affinity for video game songs.

Country	Average Intraregional affinity	Average Interregional affinity
Greece	0.858	0.756
Portugal	0.843	0.751
Germany	0.840	0.800
Malaysia	0.835	0.800
Ukraine	0.806	0.738
The Netherlands	0.804	0.788
Poland	0.802	0.740
Czech Republic	0.802	0.741
Singapore	0.801	0.783
Italia	0.789	0.775
Russia	0.787	0.787
Qatar	0.786	0.703
Indonesia	0.783	0.773
Guam	0.781	0.707
United States	0.781	0.759
Austria	0.778	0.713
Belgium	0.778	0.772
France	0.777	0.779
Costa Rica	0.774	0.690
Spain	0.773	0.768
Denmark	0.772	0.756
Israel	0.771	0.715
Norway	0.771	0.775
Egypt	0.764	0.654
Finland	0.761	0.696
New Zealand	0.757	0.751
Canada	0.754	0.739
Venezuela	0.747	0.664
Kuwait	0.743	0.693
Philippines	0.741	0.733
Brazil	0.730	0.753
Lebanon	0.729	0.682
Argentina	0.726	0.667
Colombia	0.725	0.695
United Kingdom	0.719	0.725
Australia	0.719	0.721
Brunei Darussalam	0.717	0.664
Sweden	0.716	0.751
Mexico	0.715	0.742
Puerto Rico	0.709	0.680
Kazakhstan	0.704	0.664
Turkey	0.703	0.663
United Arab Emirates	0.698	0.690
Ireland	0.694	0.681
Vietnam	0.683	0.670
Chile	0.679	0.671
Hong Kong	0.678	0.667
Romania	0.665	0.652
Japan	0.654	0.629
Switzerland	0.643	0.645
Ecuador	0.631	0.663
South Korea	0.627	0.634
South Africa	0.624	0.644
Saudi Arabia	0.623	0.633
India	0.605	0.621
Thailand	0.587	0.622
China	0.554	0.493
Taiwan	0.444	0.547

Table D.5: Intraregional affinity for pop songs.

Country	Average Intraregional affinity	Average Interregional affinity
Denmark	0.748	0.568
Norway	0.668	0.496
Belgium	0.633	0.483
Indonesia	0.613	0.442
Romania	0.599	0.414
Ukraine	0.585	0.447
Finland	0.569	0.442
Philippines	0.567	0.420
Singapore	0.558	0.449
Portugal	0.547	0.424
Greece	0.535	0.435
The Netherlands	0.526	0.421
Poland	0.513	0.416
Ireland	0.510	0.413
Venezuela	0.506	0.405
Austria	0.505	0.431
Ecuador	0.501	0.394
Sweden	0.488	0.411
Jordan	0.488	0.390
Canada	0.487	0.390
Spain	0.482	0.402
Colombia	0.482	0.419
UK	0.474	0.385
Czech Republic	0.472	0.383
Italy	0.472	0.402
Chile	0.465	0.385
USA	0.455	0.378
Argentina	0.448	0.381
Kazakhstan	0.420	0.323
Malaysia	0.412	0.362
Russia	0.411	0.375
Australia	0.407	0.358
Hong Kong	0.406	0.367
Iran	0.398	0.371
South Korea	0.398	0.371
Thailand	0.396	0.363
Brazil	0.379	0.346
New Zealand	0.378	0.361
Lebanon	0.367	0.350
Germany	0.362	0.338
Mexico	0.353	0.334
Vietnam	0.350	0.300
Puerto Rico	0.348	0.344
Costa Rica	0.339	0.321
Turkey	0.329	0.341
Switzerland	0.297	0.332
Japan	0.283	0.315
France	0.280	0.296
Taiwan	0.263	0.290
Egypt	0.244	0.291
South Africa	0.228	0.281
Bahrain	0.223	0.256
Iraq	0.222	0.227
Israel	0.168	0.224
India	0.164	0.237
United Arab Emirates	0.154	0.203
Qatar	0.141	0.201
Kuwait	0.119	0.192
China	0.109	0.176
Saudi Arabia	0.084	0.161

TABLE D.6: Intraregional affinity for folk songs.

Works Cited

Aarden, Bret and David Huron. "Mapping European Folksong: Geographical Localization of Musical Features." *Computing in Musicology* 12 (2001): 169–183. Print.

Agawu, Kofi. "Representing African Music." *Critical Inquiry* 18.2 (Winter 1992): 245–266. Print.

Analytics, Strategy. "Strategy Analytics: Worldwide Smartphone Population Tops 1 Billion in Q3 2012." Oct. 2012. Web. 17 Oct. 2012.

Barth, George. *The Pianist as Orator: Beethoven and the Transformation of Keyboard Style*. Cornell University Press, Jan. 1993. Print.

Bartók, Bela. "Race Purity in Music." *Modern Music* xix.3 (Mar. 1942): 403–406. Print.

Bengtsson, Ingmar and Alf Gabrielsson. "Analysis and Synthesis of Musical Rhythm." *Studies of Musical Performance*. Ed. Johan Sundberg. Stockholm: Royal Swedish Academy of Music, 1983. 27–60. Print.

Chafe, Chris. "Statistical Pattern Recognition for Prediction of Solo Piano Performance." *Proc. Intl. Computer Music Conf. Thessaloniki [ICMC 1997]* (Sept. 1997): 145–148. Print.

Das, M., David M. Howard, and Stephen L. Smith. "Motion Curves in Music: The Statistical Analysis of Midi Data." *EuroMicro* 2 (1999): 2013–2019. Print.

Drake, Carolyn and Caroline Palmer. "Skill acquisition in music performance: relations between planning and temporal control acquisition of skills." *Cognition* 74 (2000): 1–32. Print.

Elwin, Verrier and Shamrao Hivale. "Folk-songs of the Maikal Hills." *London, Etc.* (1944). Print.

Fallows, David. "Tempo and expression marks." *Grove Music Online. Oxford Music Online*. 22 May 2013. Web. 08 Aug. 2013.

Fost, Dan. "Cinebeat App by Smule makes music videos." May 2013. Web. 10 May 2013.

Friberg, Anders and Johan Sundberg. "Does music performance allude to locomotion? A model of final ritardandi derived from measurements of stopping runners." *Journal of the Acoustical Society of America* 105.3 (Mar. 1999): 1469–1484. Print.

Grove. "Espressivo." Aug. 2013. Web. 11 Aug. 2013.

Hamilton, Robert, Jeffrey Smith, and Ge Wang. "Social Composition: Musical Data Systems for Expressive Mobile Music." *Leonardo Music Journal* 21 (2011): 57–64. Print.

Helmholtz, Hermann L. *On the Sensations of Tone as a Physiological Basis for a Theory of Music*. Vol. 2nd English ed. of 3rd German ed (1st ed. was 1863). Dover, New York, 1954. Print.

IDC. "Android and iOS Combine for 92.3 percent of All Smartphone Operating System Shipments in the First Quarter While Windows Phone Leapfrogs BlackBerry, According to IDC." May 2013. Web. 16 May 2013.

Ives, Charles. *Essays before a Sonata, The Majority, and Other Writings*. Ed. Howard Boatwright. Norton, 1999. Print.

Knösche, Thomas R., et al. "Perception of phrase structure in music." *Hum Brain Mapp* 24.4 (Apr. 2005): 259–73. Print.

Kolisch, Rudolph. "Tempo and Character in Beethoven's Music." *MQ* xxix (1943): 169–87, 291–312. Print.

Palmer, Caroline. "Music Performance." *Annual Review of Psychology* 48 (Feb. 1997): 115–138. Print.

Repp, Bruno H. "Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's 'Träumerei'." *Acoust. Soc. Am.* (1992): 2546–2568. Print.

Rodgers, Joseph Lee and W. Alan Nicewander. "Thirteen Ways to Look at the Correlation Coefficient." *The American Statistician* 42.1 (Feb. 1988): 59–66. Print.

Sachs, Curt. *The Wellsprings of Music*. The Hague: Martinus Nijhoff, 1962. Print.

Sadakata, Makiko. "Ritme and Rizumu, Studies in music cognition." Diss. RU Radboud Universiteit Nijmegen, 2006. Print.

Sapp, Craig Stuart. "Comparative Analysis of Multiple Musical Performances." *ISMIR Proceedings* (2007): 497–500. Print.

—."Hybrid Numeric/Rank Similarity Metrics for Musical Performance Analysis." *ISMIR Proceedings* (2008): 501–506. Print.

Schaffrath, H. and D. Huron. *The Essen Folksong Collection in the humdrum kern format*. Ed. D. Huron. Center for Computer Assisted Research in Humanities, 1995. Print.

Schaffrath, Helmut. *Datenbank zur Erfassung von Musik auf Tonträgern*. Vol. 2. Bern: Peter Lang, 1987. 183–197. Print.

Seashore, Carl E. *In Search of Beauty in Music: A scientific approach to musical esthetics*. New York: Ronald Press, 1947. Print.

—."The Psychology of Music." *Music Educators Journal* 25.4 (Feb. 1939): 78. Print.

Shanahan, Daniel and David Huron. "Interval Size and Phrase Position: A Comparison between German and Chinese Folksongs." *Empirical Musicology Review* 6.4 (Oct. 2011): 187–197. Print.

Shostakovich, Dmitri. *Testimony: The Memoirs of Dmitri Shostakovich*. Ed. Solomon Volkov. Limelight Editions, 1979. Print.

Sridharan, Devarajan, et al. "Neural Dynamics of Event Segmentation in Music: Converging Evidence for Dissociable Ventral and Dorsal Networks." *Neuron* 55.3 (Aug. 2007): 521–532. Print.

Stevens, Halsey. *The Life and Music of Béla Bartók*. Clarendon Press, Oxford, 2002. Print.

Taruskin, Richard. *Oxford History of Western Music*. Vol. 4. Oxford University Press, 2005. Print.

Thiemel, Matthias. "Agogic." *Grove Music Online. Oxford Music Online. Oxford University Press. Web.* (Nov. 2013). Web.

Todd, Neil P. "A Model of Expressive Timing in Tonal Music." *Music Perception: An Interdisciplinary Journal* 3.1 (Fall 1985): 33–57. Print.

——."The Dynamics of Dynamics: A Model of musical expression." *J. Acoust. Soc Am* 91.6 (1992): 3540–3550. Print.

Truslit, Alexander. "Gestaltung und Bewengung in der Musik, as interpreted by Repp BH (1993)." *Haskins Laboratories Status Report on Speech Research* SR.111/112 (1992): 265–271. Print.

Wang, Ge. "The World Is Your Stage: Making Music on the iPhone." *Oxford Handbook of Mobile Music Studies*. Ed. S. Gopinath and J. Stanyek. Oxford University Press, 2013. Print.

Wikipedia. "Australian Immigration Trends." Oct. 2013. Web. 20 Oct. 2013.

——."Correlation Significance." Oct. 2013. Web. 09 Oct. 2013.

——."Twelve Variations on "Ah vous dirai-je, Maman"." Nov. 2013. Web. 01 Nov. 2011.