MUSICAL EVOLUTION AND HUMAN MIGRATION

MUSICAL EVOLUTION AND HUMAN MIGRATION: CLASSIFICATION, QUANTIFICATION, AND APPLICATION

By

PATRICK E. SAVAGE, B.A.(Hons)

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Master of Science

MASTER OF SCIENCE (2011) McMaster University
(Psychology, Neuroscience and Behaviour) Hamilton, Ontario

TITLE: Musical evolution and human migration: Classification, quantification, and application

AUTHOR: Patrick E. Savage, B.A. (Hons)

SUPERVISOR: Dr. Steven Brown

NUMBER OF PAGES: vi, 76

Abstract

Musical evolution and human migration: Classification, quantification, and application

Patrick E. Savage

Master of Science

Depratment of Psychology, Neuroscience and Behaviour

McMaster University

2011

The cross-cultural study of music is important to our understanding of the evolution of human biological and cultural diversity. Early comparative musicologists failed to develop rigorous scientific methods for studying this, and the modern-day fields of music cognition and ethnomusicology still lack such methods. In this thesis, I describe our attempts to design new methods for classifying and quantifying cross-cultural musical diversity and to apply these methods to the study of musical evolution and migration. Using a new method of classifying songs, we analyzed 421 songs from 16 indigenous tribes in Taiwan and the Philippines. We found striking parallels between musical and genetic diversity, both in the degree of diversity found within each culture and in the patterns of similarities between cultures. These findings suggest that music may be subject to similar processes of evolution and migration as are genes. A new, multidisciplinary, and scientifically-grounded comparative musicology may thus provide a new line of evidence to complement and integrate existing research into the complex relationship between music, biology, and culture.

Preface

This thesis presents a great deal of collaborative work that spans several disciplines and several continents. Chapters 1 and 5 were written solely by me, but Chapters 2-4 (the "meat" of my "sandwich thesis") were written collaboratively with my co-authors.

Chapter 2 was led entirely by me. In addition to writing the initial draft, designing the classification scheme with my supervisor, Steven Brown (SB), classifying the songs, and analyzing the data, I also trained and supervised Emily Merritt (EM), an undergraduate research volunteer, in order to collect the data on classification reliability. Chapter 3 was led by Tom Rzeszutek (TR), who wrote the initial draft and analyzed the data. I coded the musical data and led the creation of the musical database in consultation with SB, TR, Ying-fen Wang (YW) and Victor Grauer (VG). Chapter 4 was initially drafted by me, but contained both my own musical codings and those of VG. All data analysis for Chapter 4 was performed by TR, who was also in charge of overseeing the genetic database consisting of data supplied by Jean Trejaut (JT), Marie Lin (ML), Albert Ko (AK), Fred Delfin (FD) and Mark Stoneking (MS). SB conceived of the initial ideas for the project as a whole and set up our relationships with our collaborators. Since TR and I joined the lab in 2009, we have been equally involved with SB in all conceptual aspects of all study design and manuscript revision.

I would like to acknowledge my gratitude for all the support I have been given through the past two years. None of it would have been possible without the vision of my supervisor and friend, Steven Brown, and the incredible amount of time and money he's dedicated to training me, sending me off to conferences, helping me rehearse talks and revise manuscripts, putting up with endless bizarre ideas and email rants, and even helping to find my wife a job and giving us a place to stay while we looked for an apartment. Tom Rzeszutek has also taught me so much and likewise put up with endless ramblings, arcane discussions about the fine points of single-coding vs. multi-coding, and requests for every possible permutation of different analyses. Steven, Tom, Michel Belyk, the rest of the NeuroArts lab and the entire McMaster University Psychology, Neuroscience & Behaviour Department have provided a wonderful mix of friendship, discussion, fun, and softball that have kept my mind, body and soul nourished through these past two years. I am also grateful to my committee members, Laurel Trainor, Jonathan Stone, and to our numerous manuscript readers and collaborators for their willingness to share their time, advice, and data. Although I have not met him, I am supremely grateful to the late Alan Lomax, whose vision and passion for celebrating the beauty and diversity of the world's music is the foundation upon which our research is built. Most important of all, I am grateful to my wife, Sawa, my brother, Kelly, and my parents, Mike and Martha, for their love, understanding and support in allowing me to follow my passions for music and science.

This research was funded by the McMaster University Department of Psychology, Neuroscience, and Behaviour, my Amherst College Roland Wood Fellowship, and a grant to Steven Brown from the Social Sciences and Humanities Research Council of Canada (SSHRC).

Contents

List of figures and tables

Chapter 1

Introduction: Toward a new comparative musicology

Patrick E. Savage

It may seem strange to find a thesis focused on comparative analyses of Taiwanese aboriginal music coming out of the McMaster University Department of Psychology, Neuroscience & Behaviour. Stranger still, this thesis involves neither the perceptual experiments that are a hallmark of contemporary music cognition, nor the ethnographic fieldwork that is a hallmark of contemporary ethnomusicology. This might cause some to accuse this thesis of simply re-enacting the same mistakes found in the "armchair ethnomusicology" of early comparative musicologists, who often attempted to make conclusions about world musical diversity without first-hand experience of the cultural context in which their recordings were created. However, I believe that there is a need for a new, multidisciplinary, comparative musicology that bridges the divide between music cognition and ethnomusicology that has been the focus of recent debate (e.g., Becker, 2009a, 2009b, and responses). This thesis outlines some of the first steps towards this new comparative musicology.

In some ways, this "new comparative musicology" is not really new, because it has essentially the same goals as outlined by Adler (1885), who first defined both comparative musicology and musicology in general. What makes it new is its methodologies. For one thing, this thesis highlights the methodological importance of collaboration with experts in other cultures and other fields (something rarely seen in both early comparative musicology and contemporary ethnomusicology) to increase breadth without greatly sacrificing depth.

For another thing, this thesis highlights the importance of testing qualitative theories with rigorous quantitative methods. The chief criticism of early comparative musicologists was that their methodologies were premised on biased assumptions of European superiority that were widespread at the time (Rehding, 2000). Of course, we are all products of our times, and it is impossible to ever completely guard against biased interpretation. However, quantitative methods (particularly when accompanied by graphs and other visual aids) often make it easier to evaluate competing theories and harder for authors to get away with unsupported assertions. For example, in Ch. 3 our quantitative approach allows us to clearly reject Lomax's (1968) claim of "stylistic homogeneity" in favour of the competing claim of high internal musical diversity within cultures (Henry, 1976).

There were four specific areas in which I felt that more rigorous methods were especially critical to a revival of comparative approaches. Methodological problems with these areas have long been debated by ethnomusicologists, and resurfaced recently in two special issues of *The World of Music* devoted to Grauer's (2006) claim that today's music contains "echoes of our forgotten ancestors" that can be traced back tens of thousands of years to the African origins of anatomically modern humans. These areas are also often cited as being impediments to a better cross-cultural understanding of music cognition:

1) **Universals** (which, if any, musical features are shared by most cultures?)

2) **Classification** (how can we describe musical similarities and differences in a cross-culturally appropriate way?)

3) **Quantification** (how can we quantitatively compare these similarities and differences both between songs and between cultures?)

4) **Evolution/migration** (how did contemporary musical diversity arise?)

My supervisor, Steven Brown , has recently addressed the first area of **universals** (Brown & Jordania, in press), leaving the remaining three areas for my fellow graduate student, Tom Rzeszutek , and me to tackle for our theses. The meat of this thesis (Chapters 2-4) consists of three manuscripts addressing these areas that we are preparing to submit to academic journals. **Classification** is addressed in Chapter 2, which we plan to submit to *Analytic Approaches to World Music*. Hhere we outline our new cross-cultural song classification scheme that seeks to improve the reliability of older schemes. **Quantification** is addressed in Chapter 3, which we plan to submit to *Proceedings of the Royal Society B: Biological Sciences*. Here we outline our new methods for quantifying musical diversity both within and between cultures. **Evolution/migration** is addressed in Chapter 4, where we apply the methods in Chapters 2 and 3 to examining correlations between musical and genetic diversity among the aboriginal tribes of Taiwan. Due to the curious discrepancies between parallel genetic samples described in that chapter, we are waiting until the new genetic data is published to do final revisions on that manuscript and decide on an appropriate journal for publication. Therefore, that chapter, like Chapters 1 and 5, is formatted in standard APA style, while Chapters 2 and 3 are formatted according to the journals' specifications.

REFERENCES

Adler, G. (1885). The scope, method and aim of musicology [in German]. *Vierteljahrsschrift fur Musikwissenschaft*, *1*, 5-20.

Becker, J. (2009a). Ethnomusicology and empiricism in the twenty-first century. *Ethnomusicology*, *53*(3), 479-501.

Becker, J. (2009b). Crossing boundaries: An introductory essay. *Empirical Musicology Review*, *4*(2), 45–48.

Brown, S., & Jordania, J. (in press). Universals in the world's music. *Psychology of Music*.

Grauer, V. (2006). Echoes of our forgotten ancestors. *The World of Music*, *48*(2), 5-59.

Henry, E. O. (1976). The variety of music in a North Indian village: Reassessing Cantometrics. *Ethnomusicology*, *20*(1), 49-66.

Lomax, A. (Ed.). (1968). *Folk song style and culture*. Washington, DC: American Association for the Advancement of Science.

Rehding, A. (2000). The quest for the origins of music in Germany circa 1900. *Journal of the American Musicological Society*, *53*(2), 345-385.

Chapter 2

CantoCore: A New Cross-Cultural Song Classification Scheme

Patrick E. Savage, Emily Merritt, Tom Rzeszutek and Steven Brown

Department of Psychology, Neuroscience and Behaviour, McMaster University, Hamilton, ON, Canada

Correspondence to:

Patrick Savage

Department of Psychology, Neuroscience and Behaviour

McMaster University

1280 Main St. West

Hamilton, ON, L8S 4M9

Canada

Phone: 1 (905) 525 9140 x21443

e-mail: savagepe@mcmaster.ca

Classification of organisms and languages has long provided the foundation for studying biological and cultural history, but there is still no accepted scheme for classifying songs cross-culturally. The best candidate, Lomax and Grauer's "Cantometrics", did not spawn a large following due, in part, to concerns about its reliability. We present here a new classification scheme, called "CantoCore", that is inspired by Cantometrics but emphasizes its "core" structural characters rather than more subjective characters of performance style. Using both schemes to classify the 30 songs from the Cantometrics Consensus Tape, CantoCore appeared to be approximately 80% more reliable than Cantometrics. Nevertheless, Cantometrics still demonstrated significant reliability for all but its instrumental characters. Future multidisciplinary applications of CantoCore and Cantometrics to the cross-cultural study of musical similarity, musical evolution, musical universals, and the relationship between music and culture will provide the true test of each scheme's value.

INTRODUCTION

Musical classification is a topic that has received scant attention since the heyday of comparative musicology during the first half of the 20th century. Fields like biology and linguistics have long relied on classification as the starting point for developing broader theories, such as Darwin's (1859) theory of evolution and Jones' (1807) theory of prehistoric connections among speakers of Indo-European languages. Today, global linguistic classification databases such as the *Ethnologue* (Lewis 2009) and the *World Atlas of Language Structures* (Haspelmath et al. 2005) are fundamental to the study of language evolution, linguistic universals, and human history (Currie and Mace 2009; Dunn et al. 2011; Atkinson 2011). Musicology, in contrast, never entered into a comfortable relationship with cross-cultural classification, despite early attempts in that direction (Hornbostel and Sachs 1914; Lomax 1968). Even global music collections like the *Garland Encyclopedia of World Music* (Nettl et al. 1998) and *Smithsonian Global Sound* (http://glmu.alexanderstreet.com) that are organized according to geographic and ethnolinguistic classifications do not use an explicitly *musical* classificatory framework.

A consideration of the historical roots of the field shows that classification was central to the first definition of comparative musicology:

> *...comparative musicology has as its task the comparison of the musical works – especially the folksongs – of the various peoples of the earth for ethnographical purposes, and the classification of them according to their various forms.* (Adler 1885:14)

Although classification, comparison, and ethnography were all equal parts of this original definition, the field later changed its name to "ethnomusicology" and developed a methodological emphasis on single-culture ethnography over cross-cultural classification and comparison. This was part of a broader trend in anthropology in the wake of World War II toward cultural relativism and away from universalism (Geertz 1973). One outcome of this shift was the recognition of a theoretical distinction between "etic" (objective, outsider) and "emic" (subjective, insider) theories of classification (Harris 1976). This dichotomy nicely characterizes the paradigmatic difference between early comparative musicology and contemporary ethnomusicology. Ethnomusicologists have largely rejected etic, acoustic classification schemes, despite a plea for pluralism in approaches to world musics (Merriam 1982; Nettl 2005; Agawu 2010). Although the goal of classifying musics acoustically presents many challenges – for example, the need for classification schemes to be universally applicable – these challenges do not *a priori* invalidate cross-cultural classification (but see Hood 1971; Blacking 1973; McLeod 1974).

Along these lines, are two major methodological challenges to classifying music cross-culturally. One challenge is specific to instrumental music: how do we ensure that we are comparing like with like when different cultures use different instruments with differing acoustic features, production mechanisms, and tuning systems (Ellis 1885)? The second is specific to vocal music: how can we design a classification system that is broad enough to accommodate all musical cultures while maintaining a distinction between "song" and "speech"? While the instrumental classification scheme of Hornbostel and

Sachs (1914) is still widely used today, there remains no widely accepted song-classification scheme.

One solution to the problem of song classification is to see the relationship between music and language as a continuum – a "musilinguistic" spectrum (Brown 2000) – rather than as a contrast between two discrete domains. This acknowledges the general consensus in ethnomusicology, neuroscience, and evolutionary biology that music and language are deeply interconnected (Darwin 1871; Feld and Fox 1994; Wallin, Merker, and Brown 2000; Patel 2008). The implications for song classification are clear: a truly universal approach cannot exclude "non-musical" vocalizations but must accommodate any type of vocalization sitting along the musilinguistic spectrum of communicative forms from speech, to songs, to everything in between. While Sachs (1943) proposed such a spectrum in his distinction between "logogenic" (word-born) and "melogenic" (melody-born) songs, there is a need for a classification scheme that can accommodate the diversity of ways in which song-features can independently vary across *multiple* musilinguistic spectra. For example, some songs can have irregular "speech-like" (parlando) rhythms but use discrete "music-like" pitches, while others can have metric "music-like" rhythms but use indeterminate "speech-like" pitches.

Multi-dimensional, musilinguistic spectra are in fact a major design feature of the best-established song-classification scheme to date, "Cantometrics" (Lomax and Grauer 1968; Lomax 1976). Cantometrics classifies songs according to 37 acoustic characters related to their structure, performance style, and instrumental accompaniment. Each character contains between 3 and 13 character-states, which are ordered along a social continuum from "individualized" to "groupy". This continuum can be thought of equally well as a musilinguistic continuum, since speech tends to be more individual-oriented and song more group-oriented.

Applying this scheme to a global sample of thousands of songs from hundreds of cultures, Lomax found that global song diversity was organized into 10 major stylistic families that also correlated with extra-musical features of social structure and historical contact. Critics generally applauded this ground-breaking attempt to quantitatively address the relationship between music and culture and supported its broad findings, despite some concerns over methodological issues regarding sampling, treatment of intra-cultural diversity, and the interpretation of correlations between music and social structure (Naroll 1969; Driver 1970; Downey 1970; Nettl 1970; Maranda 1970; Henry 1976; Erickson 1976; Dowling and Harwood 1986; Grauer 2005; Leroi and Swire 2006). However, many critics were divided over Lomax's emphasis on performance style over song structure. Lomax's agenda in creating Cantometrics was to replace Western musicology's traditional emphasis on musical structure and notation – which he and many others saw as being Eurocentric and elitist (Lomax 1959; Feld and Fox 1994) – with a more performance-oriented system. While some critics supported the development of measurements of performance characters such as "nasality" and "rasp", others were concerned that such characters were overly subjective and thus unreliable (Downey 1970; Maranda 1970).

The principal objective of the current study is to present a detailed analysis of a new universal song-classification scheme. We call it "CantoCore" because of its emphasis on the "core" structural characters of song. The scheme takes its lead from the updated 1976 version of Cantometrics but focuses only on characters of song-structure rather than performance-style or instrumentation (see Fig. 1), because of our prediction that structural characters should be more reliable. We have reorganized, supplemented, and attempted to more objectively operationalize these characters, building on the work of others when possible (Kolinski 1961, 1962, 1973; Plomp and Levelt 1965; Patel and Daniele 2003; Leroi and Swire 2006; Busby 2006). In addition, the scheme introduces several structural characters not present in Cantometrics, most notably those related to scales and rhythms. Finally, the scheme is designed to accommodate musical forms at all points along the musilinguistic spectrum, from a simple sentence to the most complexly-textured responsorial polyphony. The current study also includes a test of the inter-rater reliability of song codings, comparing 1) CantoCore vs. Cantometrics, and 2) the structural characters of Cantometrics vs. its performance and instrumental characters. To accomplish this, we use the global set of 30 songs contained in the Cantometrics Consensus Tape (Lomax 1976) that Lomax selected to demonstrate the cross-cultural validity of the Cantometrics scheme.

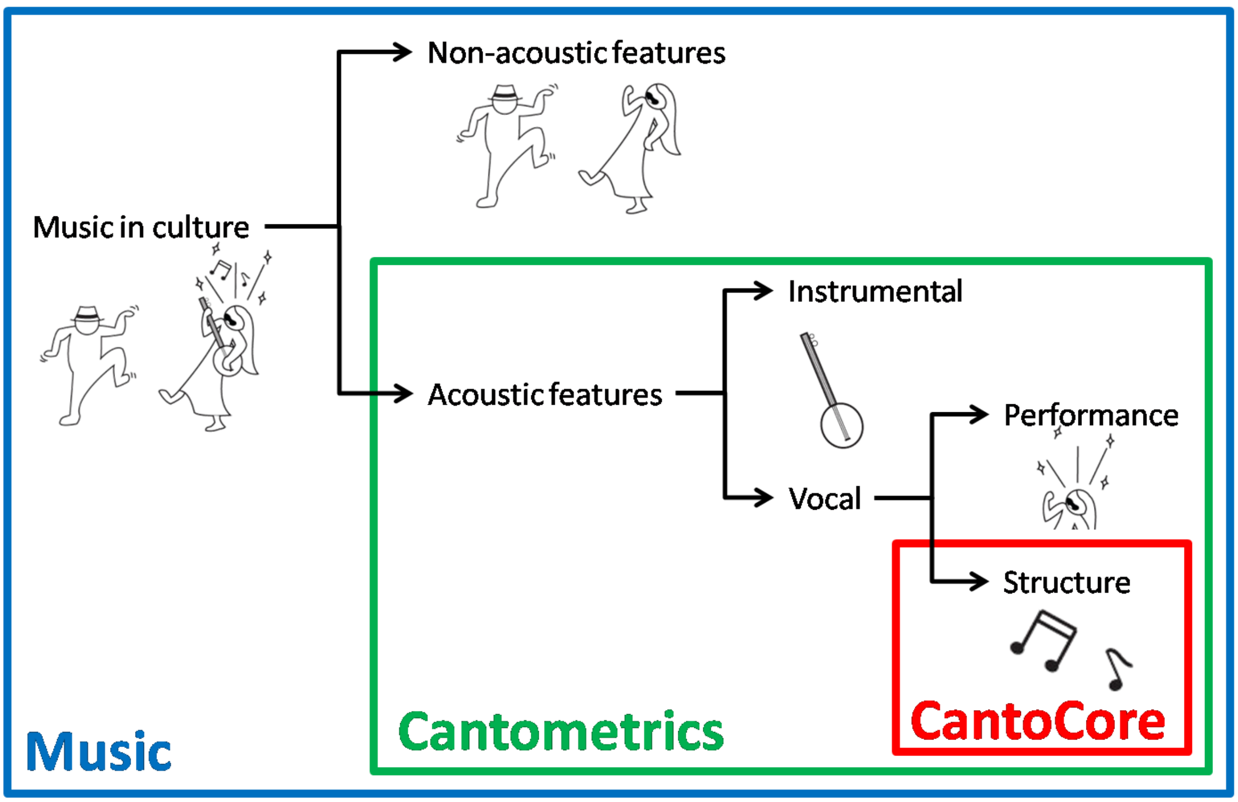


Fig. 1: A comparison of the types of musical characters classified by CantoCore vs. Cantometrics. Both classification schemes rely exclusively on acoustic information rather than on non-acoustic characters. But whereas Cantometrics (green box) focuses on both the performance and structural characters of songs as well as their instrumental accompaniment, CantoCore (red box) focuses exclusively on the structural characters of the vocal part, excluding both performance and instrumental characters.

CLASSIFICATION SCHEME

Theoretical framework

The musical hierarchy. Music is a hierarchical system made up of several levels of organization (Schenker 1979; Lerdahl and Jackendoff 1983; Krumhansl 1990; Anku 2000; Tenzer 2006). Figure 2a presents a schematization of the musical hierarchy that we employ in organizing the characters of the CantoCore classification scheme, themselves listed in Figure 2b. A useful analogy for conceptualizing our classification scheme is to think of a song as a biological organism. In essence, songs are simply a complex combination of notes, just as organisms are a complex combination of cells. However, as with the cells in an organism, the notes in a song interact with each other and with their extra-musical environment at many different levels and in many different ways. These complex interactions can never be fully quantified but can still be usefully modeled.

The most basic distinction is that between the *note* level – where the note is regarded as the basic building block of music – and the *supra-note* level. The note level consists of three characters: 1) *rhythm* (colored red in Figure 2a), reflecting the relative duration of a note; 2) *pitch* (blue), reflecting the acoustic frequency of a note; and 3) *syllable* (green), reflecting the articulatory configuration of a sung note (exemplified by "la" in the figure). The supra-note domain consists of interactions between notes, as organized into three broad hierarchical domains: 1) *phrase*, representing the between-note level within individual vocal parts, 2) *texture*, representing the between-part level, in which simultaneous phrases in different vocal parts overlap in time, and 3) *form*, representing the between-phrase level, where successive phrases combine to form larger melodic units. Figure 2b lists the classification characters associated with each of these three supra-note domains. It also shows that the domain of "phrase" contains the three note-level characters of rhythm, pitch, and syllable (color coded the same as in Fig. 2a). CantoCore classifies 26 structural characters of songs (Fig. 2b), organized into categories associated with the note and supra-note domains listed above. Fifteen of these characters are refined versions of structural characters already contained in Cantometrics, while 11 characters – mostly those related to rhythm and scale – are new, as indicated by asterisks in the detailed scheme below.

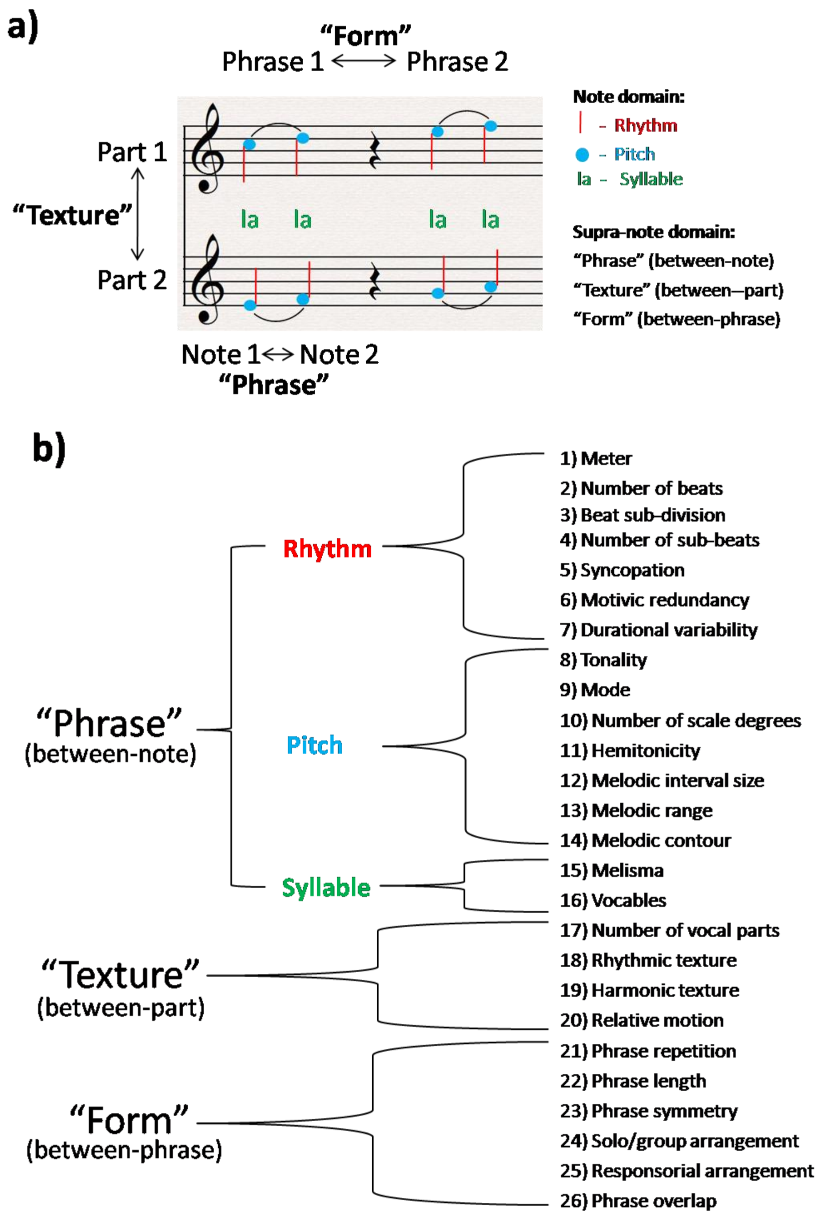


Fig. 2: a) The musical hierarchy is comprised of "note" and "supra-note" domains. The three main note domains are rhythm (red), pitch (blue), and syllable (green), as represented by the sung note "la". Interactions between notes give rise to the supra-note domains of "phrase" (the between-note level), "texture" (the between-part level) and "form" (the between-phrase level). b) The 26 structural characters that comprise the CantoCore classification scheme are organized according to these note and supra-note domains.

Quantitative vs. qualitative characters. A fundamental distinction in classification theory is that between *quantitative* (or continuous) characters and *qualitative* (or discrete) characters (Sneath and Sokal 1973). Quantitative traits can be classified with regard to their size. For example, melodic intervals (character **12** in the CantoCore scheme) vary in a continuous manner from very small intervals to very large, and everything in between. Another way to code characters quantitatively is with regard to their frequency of occurrence in a song. In CantoCore, vocables (character **16**) are coded with regard to their frequency of occurrence, ranging from being completely absent (low frequency) to being ubiquitous (high frequency). Qualitative traits, by contrast, cannot be placed onto a numerical spectrum of size or frequency, and are instead organized as a series of discrete states. For example, melodic contours (character **14**) come in a variety of discrete types, such as descending contours, ascending contours, arched contours, and the like. Of the 26 CantoCore characters, 15 are quantitative traits and 11 are qualitative traits by the standards of classification theory.

Ordering of character-states. Most of CantoCore's 26 characters are divided into 3-4 character-states, resulting in a total of 97 character-states across the scheme. Of these, 55 are new to the scheme, as indicated by asterisks in the detailed description below. Figure 3 represents our rationale for ordering the character-states within each character. Character-states are ordered in a consistent manner, spanning a musilinguistic spectrum from language-like (left side) to music-like (right side). However, the method for achieving this differs for quantitative and qualitative characters, as shown in Figure 3 above and below the horizontal arrow.

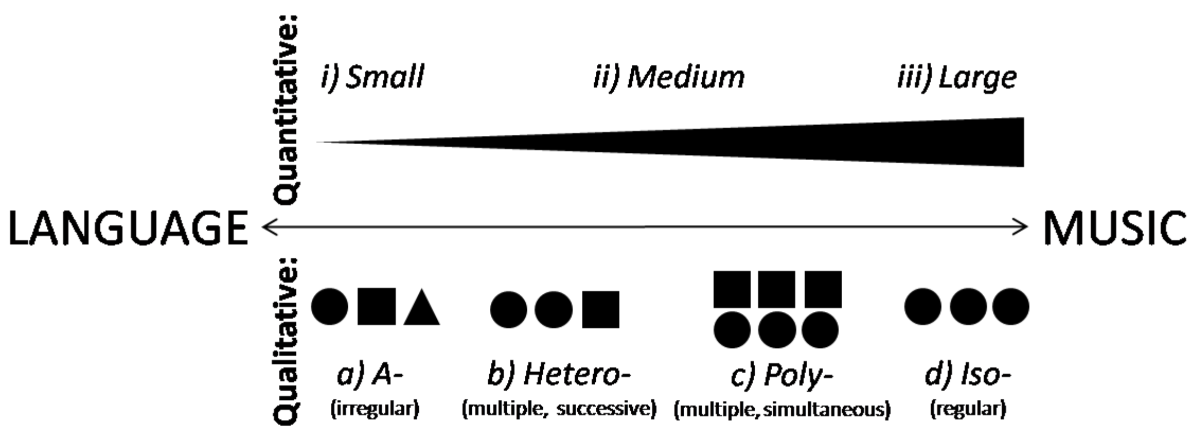


Fig. 3: The character-states within each character are organized according to a "musilinguistic" spectrum spanning from language-like to music-like (no value judgment is implied). Quantitative characters (top part of the figure) are ordered in terms of increasing size from small to large using lower-case roman numerals. Qualitative characters (bottom part of the figure) are ordered in terms of increasing "regularity" using lower-case letters from irregular ("A-") to regular ("Iso-"), with semi-regular states between them having either multiple successive forms ("Hetero-") or multiple simultaneous forms ("Poly-"). The geometric shapes are used for heuristic purposes only to demonstrate the various facets of regularity.

For quantitative characters, character-states are listed in order of increasing size or frequency using lower-case roman numerals (i, ii, iii, etc). This allows for precise placement of states along a continuum spanning from small (speech-like) to large (song-like). For qualitative characters, character-states are listed in order of increasing "regularity" using lower-case letters (a, b, c, etc.), spanning from irregular (speech-like) to regular (song-like). By regularity, we refer to the degree of repetitiveness of a character throughout a song, where redundancy is far more associated with music than speech (Lomax 1968). Because qualitative characters could not always be divided up consistently, we employed a series of prefixes to convey a spectrum of qualitative states (see the geometric shapes at the bottom half of Fig. 3 as a guide): a) “A-“ implies that a feature is absent from a song; b) “Hetero-” implies that multiple but successive features occur; c) “Poly-” implies that multiple simultaneous features occur; and d) “Iso-” implies that a single feature occurs consistently throughout a song. Applying these concepts to meter, for example, we can see that irregular “a-metric” songs have no discernable meter; semi-regular “hetero-metric” or “poly-metric” songs have multiple meters that are present successively or simultaneously, respectively; and regular “iso-metric” songs have a single, constant meter throughout.

Classification logistics

Within-song heterogeneity. Reality is too complex to be fully captured in a single classification. Songs change over time and can contain multiple sections whose codings conflict with one another. Some important work has been done regarding quantifying this kind of dynamic heterogeneity with regards to specific characters such as interval size and note duration (Toiviainen and Eerola 2001; Huron 2006). However, there is also a need for broader classification schemes that provide simpler classifications but span a number of characters across multiple domains.

Maximal values. Heterogeneity can be partially accommodated for quantitative characters by defining them with regard to summary statistics describing their size or frequency. Hence, a song that has multiple states for such characters could be coded with regard to things like their maximal value for that song, their mean value for the song, or their standard deviation. For consistency, and to make the scheme possible to use quickly by ear without resorting to laborious transcription and note-counting, quantitative characters have been defined in terms of maximal values and divided into the character-states of “small”, “medium”, and “large” by imposing somewhat arbitrary thresholds. This is intended to reduce the amount of theoretical expertise and time required to code the songs. If one is working from notated scores or transcriptions, or if the coder has enough confidence in his/her ability to hear very fine distinctions, the raw numerical values may be used to increase precision (see Figure 4 below). However, this may give an appearance of precision that is unrealistic, as we found that making the scheme finer-grained did not improve its reliability.

Multi-coding. For qualitative characters, heterogeneity is more difficult to classify. In some cases, the heterogeneity of a song’s characters can be accommodated by character-states that specify an intrinsic heterogeneity of features (e.g., “hetero-metric”, “poly-tonal”). However, in other cases, this can only be accomplished by “multi-coding”, in other words selecting *multiple* distinct character-states for the same song (e.g., both “descending” and “arched” contours if both types occur in a single song). As a general rule, multi-coding should be avoided if one character-state is clearly the most prominent in a song.

Character dependence. Some characters are dependent on others. For example, “a-metric” songs that have no beat (character **1**) cannot possibly have a sub-beat (character **3**). For such characters, a “n/a” character-state is included to denote something that is unclassifiable. “?” may be used instead if recording quality or other factors make it impossible to code a given character, or if the musical characters are simply too complex to specify (following Busby 2006).

Relationship to Cantometrics. For all characters that are derived from structural characters of Cantometrics, the original Cantometrics line number and corresponding character-states names from the updated version of Cantometrics (Lomax 1976) have been given. There are a few small differences between this and the version (Lomax and Grauer 1968) used to collect the original Cantometric data, but these can be easily inter-converted. Therefore, it is basically possible to convert old Cantometric codings into CantoCore codings if desired, which may be useful in re-analyzing the original Cantometric data without having to re-code each of its thousands of songs.

Instrumental application. Due to the complications listed in the introduction involved in classifying instrumental music cross-culturally, we have designed CantoCore exclusively for the purpose of classifying vocal music. Most of the classifications could also be useful for classifying instrumental music, but caution should be exercised in doing so, particularly regarding the additional constraints on sound production and intonation that are introduced by different instrument types. For example, although breathing can still be helpful in determining phrase boundaries for aerophones, it will be less useful when dealing with chordophones.

How to code. We have attempted to define all of our terms as precisely as possible so that the coder can provide precise numeric values if they are working directly from a score or transcription, or if they have a high level of listening expertise. These definitions therefore require a modest background in music theory to code well. However, since much of the world’s music is transmitted orally and is difficult and time-consuming to transcribe, we have also aimed to create our character-states such that they can be reliably identified by ear without detailed notation. Ultimately, the numeric values are simply guidelines to assist the coder in interpreting their holistic, subjective classification of the songs. Once the coder has practiced with a few dozen songs, he/she should be able to code a 3-minute song by ear in 15-20 minutes, which is comparable to the amount of time required to do so using Cantometrics (Lomax 1976).

When coding, the coder should first listen to the song once through, jotting down important notes and trying to get a sense for the different phrases that make up the song: how many there are, in what order, how long each phrase is, what scale(s) or meter(s) (if any) underlie them, etc. The instrumental accompaniment can be used if it is helpful in interpreting the correct song classification, but if there is any conflict between the vocal and the instrumental components, the coder should focus only on the vocal component. After they have listened to the song once, they should go through and attempt to classify each character in order from 1 to 26. They should then listen to the entire song again, checking the initial codings and paying particular attention to complicated or ambiguous codings. The coder can also jump forwards or backwards within the song or repeat the song as many times as necessary to arrive at a set of codings they are confident in. Any particularly noteworthy features, such as ambiguities, striking characters not classifiable, etc., should be listed in a separate “comments” column.

This same format applies regardless of the length of the song or any extra-acoustic information about the song. The definition of what constitutes a “song” varies, but in the absence of other information, it is reasonable to assume that different tracks on recordings correspond to different songs. Song classifications should be interpreted with the help of recording liner notes, music theory (both emic and etic), and all other available resources. However, the initial classification should be done blind to extra-acoustic information as much as is practically possible (i.e., without knowing what culture the song is from or how the singer(s) classify their own music). CantoCore is fundamentally an etic, acoustic classification scheme, with all of the benefits and drawbacks that this entails (Harris 1976).

Definitions. Our goal was to create a descriptive system that allows a common vocabulary for classification, not a prescriptive system that dictates how one should perceive music. Nevertheless, for such a system to be reliable, it is necessary to have standardized definitions. Since few, if any, musical terms have cross-culturally agreed-upon definitions, we have offered our own definitions for each character, as well as for several key terms (see Box 1). Definitions about complex musical categories such as “tonality” and even seemingly simpler categories such as “interval size” have been, and will continue to be, debated. Our definitions are simply operational ones that can be usefully applied cross-culturally. Even when using these definitions, some level of disagreement and ambiguity is inevitable due to perceptual differences between individuals and between cultures. We discuss some observations on agreement in the “RELIABILITY” section.

Box 1: Glossary of key terms

- *Note:* A continuous combination of one pitch and one syllable for a fixed duration. If the pitch or syllable changes or begins afresh, this constitutes a new note.
- *Vocal part:* A series of notes sung by one voice, or by several voices in unison and/or in octaves. Slight variations between voices singing basically in unison are not counted as separate parts unless the offset between parts exceeds 0.1s in time or 50 cents in pitch (see characters **18** and **19** in the scheme).
- *Phrase:* A self-contained series of notes in one or multiple vocal parts. Phrases are usually separated by breaths or long pauses, but can also be separated by more complex grouping principles. The coder should rely on their intuition in deciding what constitutes a new phrase, focusing on breaths in ambiguous cases.
- *Beat:* Fixed time interval(s) at which notes regularly recur. The beat is often sub-divided into multiple sub-beats. In cases where the distinction between a "beat" and a "sub-beat" is ambiguous, the coder should designate the beat as the unit which feels the most natural to take steps to when dancing.
- *Tonic:* The central tone(s) that seems to be the most stable in a scale. The tonic is usually either the most common note in a scale, the final note in a phrase, or both. In ambiguous cases, the coder should designate the tonic as the note that occurs most frequently as the final note in a phrase. If the tonic seems to consistently differ between phrases or between vocal parts, this should be classified as hetero- or poly-tonal, respectively (see character **8** in the scheme).
- *Scale degree:* Notes that share the same pitch class (e.g., B, Db) regardless of their absolute pitch are considered as the same scale degree (i.e., assuming octave equivalence). Because the production of vocal pitches often fluctuates by up to 100 cents from tonal targets during normal singing (Pfordresher et al. 2010), we have followed the compromise adopted by Kolinski (1961) and others of rounding pitches to the nearest 100 cents, for a maximum of 12 possible unique scale degrees. Unfortunately, as a result of this compromise, there may be some cases in which separate micro-tones are classified as a single scale degree, while in other cases normal variation of intonation may be classified as separate scale degrees.

N.B. None of these terms have a well-agreed upon cross-cultural definition. We offer these definitions to assist in developing a shared classification vocabulary that can be reliably replicated by different coders. However, we recognize that many cultures have their own emic definitions that may differ from ours, and that there are many grey areas in which the perception and interpretation of these features may vary both within and between cultures.

THE “CANTOCORE” SONG CLASSIFICATION SCHEME

NOTE: Characters and character-states marked with an asterisk are those that are new to this scheme and that are not taken from Cantometrics.

I) “PHRASE” (between-note)

A) Rhythm

1) METER (Cantometrics Line 11)
Cyclic groupings of strong and weak beats into bars

(a) *A-metric:* No consistent beat (formerly *“parlando rubato – free rhythm”*)
(b) *Hetero-metric:* There is a consistent beat, but strong and weak beats occur without a consistent pattern (formerly divided into *“one-beat rhythm”* and *“irregular meter”*)
(c) *Poly-metric*:* Multiple cyclic patterns of strong and weak beats occur simultaneously (e.g., 6/8 against 3/4) (*“simple”* and “*complex*” poly-meter from Cantometrics Line 12 have been combined and moved here)
(d) *Iso-metric:* There is a single, consistent pattern of strong and weak beats (e.g., 3/4, 6/8, 5/4, 2+2+3/8) (formerly divided into *“simple”* and *“complex”*)

N.B. See Box 1 for the definition of “beat”. Songs not classified as (d) (“*iso-metric*”) must be coded (n/a) for characters **(2-5)**. Songs that transition between metric types (e.g., an “*a-metric*” section giving way to an “*iso-metric*” section) should be multi-coded.

Comments: The “*poly-metric*” character-state was moved here from Cantometrics Line 12. The new characters **(2-4)** were created to deal with various iso-metric sub-types unclassifiable using Cantometrics. For instance, Cantometrics did not create any distinctions between 3/4, 4/4, 9/8 and 12/8 meters, although there are important regional differences in the distribution of these metric types.

2) NUMBER OF BEATS*
The number of beats in a bar

(a) *Duple:* The number of beats can be divided by 2 (e.g., 2/4, 4/4, 6/8, 12/8, 2+3/8)
(b) *Triple:* The number of beats can be divided by 3 but not by 2 (e.g., 3/4, 9/8, 2+2+3/8)
(c) *Complex:* The number of beats can only be divided by prime numbers greater than 3 (e.g., 7/4, 5/8, 2+2+3+2+3/8)
(n/a) *A-/hetero-/poly-metric:* See **(1)**

Comments: Only the number of beats is coded here, regardless of the manner in which they are sub-divided into sub-beats, which is coded in **(3)**. For example, a 2+3/8 meter is composed of two beats, one of which is divided into two sub-beats and the other of which is divided into three sub-beats.

3) BEAT SUB-DIVISION*

Division of beats into sub-beat-level metric groupings

(a) *A-divisive:* Beats are not sub-divided (e.g., a 4/4 piece containing only ♩ and 𝅗𝅥 notes)

(b) *Hetero-divisive:* Beats are sub-divided, but the number of sub-beats per beat changes (e.g., 2+2+3/8)

(c) *Iso-divisive:* Beats are sub-divided into a consistent number of sub-beats (e.g., 6/8, a 4/4 piece containing ♪ notes)

(n/a) *A-/hetero-/poly-metric:* See **(1)**

N.B. See Box 1 for the distinction between “beat” and “sub-beat”. Songs not classified as (c) (“*iso-divisive*”) must be coded (n/a) for character **(4)**.

Comments: This character was created to capture a crucial metric dimension not classified in Cantometrics. It is almost identical to **(2)**, but captures a finer level of the metrical hierarchy, and does not have a “*poly-divisive*” character-state because this would be redundant with “*poly-metric*” (see **1**).

4) NUMBER OF SUB-BEATS*

The number of sub-beats in a beat

(a) *Simple:* The number of sub-beats can be divided by 2 (e.g., ♩ beat divided into ♪ note sub-beats; includes 3/4, 4/4, etc.)

(b) *Compound:* The number of sub-beats can be divided by 3 but not by 2 (e.g., ♩. beat divided into ♪ note sub-beats; includes 6/8, 9/8, “swing”, etc.)

(c) *Complex:* The number of sub-beats can only be divided by prime numbers greater than 3 (e.g., ♩ beat divided into 5 sub-beats)

(n/a) *A-/hetero-/poly-metric or a-/hetero-/poly-divisive:* See **(1/3)**

Comments: The “c*omplex*” character-state is included for completeness, but in practice most cultures break down groupings of five or more sub-beats into smaller groupings of twos and threes, so most songs that seem to be in 5/8, 7/8, etc. will actually turn out to be “*hetero-divisive*” songs (see **3**) that are in 2+3/8, 2+2+3/8, etc.

5) SYNCOPATION*

The percentage of notes that are stressed but in a metrically unaccented position

(i) *Un-syncopated:* <5%
(ii) *Moderately syncopated: 5-20%*
(iii) *Highly syncopated:* >20%

(n/a) *A-/hetero-/poly-metric:* See **(1)**

Comments: The term “syncopated” is used instead of Kolinski’s (1973) unwieldy term “contrametric”. We also attempt to define syncopation and recognize different degrees of syncopation, both of which Kolinski did not do.

6) MOTIVIC REDUNDANCY*

The percentage of all notes that are constructed from a single recurring rhythmic pattern

(i) *Non-motivic:* <20%
(ii) *Moderately motivic: 20-50%*
(iii) *Highly motivic*: >50%

N.B. If there are multiple motives, classify based on the frequency of the most common motif.
Comments: To give an example, in Figure 4, 40 out of the 61 notes (66%) are constructed from the rhythmic pattern: ♩ ♪ ♪ ♩

7) DURATIONAL VARIABILITY*

Maximum number of different types of duration values in a song

(i) *Low durational variability:* <3 duration values (e.g., only ♪ and ♩)

(ii) *Moderate durational variability:* 3-4 duration values (e.g., ♪, ♩ and 𝅗𝅥)

(iii) *High durational variability:* >4 duration values (e.g., 𝅘𝅥𝅯, ♪, ♩, 𝅗𝅥 and 𝅝)

N.B. Dotted notes are counted as separate duration values.
Comments: This modifies Patel & Daniele’s (2003) conception of rhythmic variability to focus on *global* variability across all notes in a song, rather than only variability between each successive pair of notes.

B) Pitch

8) TONALITY*

Organization of discrete pitches around one or more tonic notes

(a) *Indeterminate a-tonal:* No discrete pitches (e.g., exclamations, heightened speech)
(b) *Discrete a-tonal:* Discrete pitches, but no tonic
(c) *Hetero-tonal:* Tonic modulates/shifts between phrases
(d) *Poly-tonal:* Multiple, simultaneous tonics in different vocal parts
(e) *Iso-tonal:* Single tonic throughout

N.B. See Box 1 for the definition of “tonic”. Songs not classified as (e) (“*iso-tonal*”) must be coded (n/a) for characters **(9-10)**.

Comment: This character was added to make Kolinski's (1961) scale-classification scheme more universally applicable, as his scheme did not recognize the fact that some songs have no tonic or have multiple tonics.

9) MODE*

Presence of scale degrees at a minor 3rd (250-350 cents) or major 3rd (350-450 cents) above the tonic

(a) *A-modal:* No 3rd present
(b) *Hetero-modal:* Both major and minor 3rd appear in separate phrases
(c) *Poly-modal:* Both major and minor 3rd appear in the same phrase
(d) *Minor iso-modal:* Minor 3rd only
(e) *Major iso-modal:* Major 3rd only
(n/a) *A-/hetero-/poly-tonal:* See **(8)**

N.B. See Box 1 for the definition of a "scale degree".

Comments: Kolinski's (1961) concept of "mode" is a nested one that can only be applied within specific scales, not as an independent character. To make the concept more independent, we have restricted this character to the common distinction between modes with minor vs. major thirds. Characters dealing with other scale degrees besides the third or with micro-tonal intonations have been avoided due to a lack of consensus about how to classify these characters.

10) NUMBER OF SCALE DEGREES*

Number of scale degrees found in the scale

(i) *Sparse scale:* <4 scale degrees

(ii) *Moderately dense scale:* 4-5 scale degrees

(iii) *Dense scale:* >5 scale degrees

(n/a) *A-/hetero-/poly-tonal:* See **(8)**

N.B. See Box 1 for the definition of a "scale degree".

Comments: The more common term "scale degree" is used to refer to pitches that share the same note name regardless of octave, rather than Kolinski's (1961) term "tint" or the alternative term "pitch class".

11) HEMITONICITY*

Percentage of melodic intervals that are semitones (50-150 cent intervals)

(i) *Anhemitonic:* <5%
(ii) *Moderately hemitonic: 5-20%*
(iii) *Highly hemitonic:* >20%

Comments: Scales are commonly described as being simply hemitonic (containing semitones) or anhemitonic (not containing semitones). However, this dichotomy fails to recognize the importance of different gradations in the frequency with which semitones are used.

12) MELODIC INTERVAL SIZE (Cantometrics Line 21)
Maximum pitch distance between successive notes within any vocal part

(i) *Small intervals:* <350 cents (i.e., minor 3rd or less; formerly divided into "*monotone*", *"narrow"*, and "*diatonic"* intervals)
(ii) *Medium intervals:* 350-750 cents (i.e., major 3rd – perfect 5th; formerly divided into "*wide*" and "*very wide*" intervals)
(iii) *Large intervals*:* >750 cents (i.e., minor 6th or greater)

N.B. Intervals between the final note of a phrase and the first note of the next phrase are not coded.

Comments: Vague definitions from Cantometrics combining interval frequency and size, such as "intervals of a half step or less are prominent (though not necessarily dominant)", were redefined solely in terms of maximum size. A new character-state was created to recognize the importance of much larger intervals such as the octave.

13) MELODIC RANGE (Cantometrics Line 20)
Maximum pitch distance between the highest and lowest notes within any vocal part

(i) *Small range:* <750 cents (i.e., perfect 5th or less)
(ii) *Medium range:* 750-1250 cents (i.e., perfect 5th – octave)
(iii) *Large range:* >1250 cents (i.e., more than an octave)

Comments: This character is essentially unchanged from the 1976 version of Cantometrics.

14) MELODIC CONTOUR (Cantometrics Line 15)
Shape resulting from all changes in interval direction within a vocal part

(a) *Horizontal*:* No ascending or descending intervals
(b) *Ascending*:* Ascending intervals only
(c) *Descending:* Descending intervals only (formerly divided into "*descending*" and "*terraced*" contours)
(d) *U-shaped*:* First descending, then ascending intervals
(e) *Arched:* First ascending, then descending intervals
(f) *Undulating:* Multiple changes of interval direction

N.B. Each phrase should be treated as having its own contour, except when there are clear "hyper-phrase" contours that connect multiple phrases. Cases where multiple contours appear in different phrases and/or different vocal parts should be multi-coded. Some discretion must be used in deciding what constitutes a change of interval direction. In general, temporary interval changes that do not affect the overall melodic shape should be ignored. Otherwise, almost all contours will end up being classified as "*undulating*", reducing the overall informativeness of the character.

Comments: Although most music tends to be descending, three character-states needed to be added to allow for horizontal, ascending, or U-shaped contours not classifiable by Cantometrics. Cantometrics Line 19 ("Position of the final tone") was removed because it was redundant with this character.

C) Syllable

15) MELISMA (Cantometrics Line 29)
Maximum number of consecutive notes without articulating a new syllable

(i) *Syllabic:* 1-2 notes
(ii) *Mildly melismatic:* 3-5 notes
(iii) *Strongly melismatic:* >5 notes

Comment: While Cantometrics defined melisma in terms of the *frequency* of melisma, this character is defined in terms of maximum *length* to be more consistent with other quantitative characters.

16) VOCABLES (Cantometrics Line 10)
The percentage of syllables containing only vowels or semi-vowels (e.g., "y", "h", "w")

(i) *Few vocables:* <20% (formerly "*little or no repetition"*)
(ii) *Some vocables:* 20-50% (formerly divided into "*some repetition"* and "*half repetition"*)
(iii) *Many vocables:* >50% (formerly divided into "*quite repetitious"* and "*extreme repetition*")

Comment: Vocables (non-lexical "nonsense" syllables) are an important character of many musics cross-culturally, but are difficult to define for one who does not speak the language (Maranda 1970). This led Lomax to change the emphasis from vocables to textual repetition, but this change of emphasis becomes confounded with phrase repetition **(21)**. This character instead uses words containing only vowels or semi-vowels as a proxy for vocables.

II) "TEXTURE" (between-part)

17) NUMBER OF VOCAL PARTS (Cantometrics Line 4)
Maximum number of simultaneous vocal parts

(i) *One-part:* 1 (formerly divided into "*solo"* and "*unison"*)
(ii) *Two-part*:* 2
(iii) *Many-part*:* >2

N.B. See Box 1 for the definition of a "vocal part". Songs classified as (i) ("*one-part*"), including both unison and solo songs, must be coded (n/a) for characters

(18-20). "*Many-part*" songs may require multi-coding for characters **(18-20)** as may "*two-part*" songs that transition between different texture types.

Comments: This character no longer distinguishes between the number of voices singing each part or their rhythmic relationship, which are now coded in **(18)** and **(24)**, respectively. Further distinctions between the numbers of parts (e.g., 3-part, 4-part, 5-part, etc.) were avoided because it proved difficult to reliably code beyond three parts.

18) RHYTHMIC TEXTURE (Cantometrics Line 12)
Temporal asynchrony in the relative onsets of different vocal parts (in seconds)

(a) *Hetero-rhythmic (heterophonic):* 0.1–1s (formerly "*rhythmic heterophony*")
(b) *Poly-rhythmic (polyphonic):* >1s (formerly divided into "*accompanying rhythm"* and "*rhythmic counterpoint"*)
(c) *Iso-rhythmic (homophonic):* **<**0.1s (formerly "*rhythmic unison*")
(n/a) *One-part (monophonic):* See **(17)**

N.B. Unison songs where multiple singers generally sing the same pitches with less than 0.1s offset are classified as "*one-part*", not "*iso-rhythmic*" (see Box 1 and **17**). Songs with different rhythmic textures between different vocal parts or different phrases should be multi-coded. Songs not classified as "*iso-rhythmic*" must be coded "n/a" for character **(19)**.

Comments: The corresponding terms "poly-/hetero-/homo-/mono-phonic" have been included because they are commonly used to categorize texture as a whole, despite ambiguities about distinguishing between rhythmic texture, harmonic texture, and relative motion.

This character concerns the rhythmic relationship between the notes of multiple parts, regardless of what meter those parts are in. Thus, "*iso-metric*" songs (see **1**) with rhythmically independent parts can be "*poly-rhythmic*" despite not being "*poly-metric*". The two types of poly-meter that were originally also classified in this character have been moved to **(1)**.

19) HARMONIC TEXTURE*
Minimum harmonic interval (octave-equalized – see N.B. below) between simultaneous vocal parts that is sustained for at least 1 second

(i) *Dissonant:* 50-249 cents (includes 951-1150 cents) (e.g., 2nds/7ths)
(ii) *Consonant:* 250-600 cents (includes 600-950 cents) (e.g., 3rds-6ths)
(n/a) *One-part (includes 0-49 and 11501-1200 cents), or poly-/hetero-rhythmic:* See Box 1 and **(17/18)**

N.B. Harmonic intervals should be calculated after correcting for absolute differences in pitches by transposing them to the octave that minimizes the harmonic interval. For example, the top note in a harmonic interval of 1000 cents (minor 7^{th}) can be transposed down one octave to create a harmonic interval of

200 cents (major 2^{nd}). Therefore, the largest possible harmonic interval is 600 cents (a tritone) before the octave-equalized interval size begins to decrease again.

Comments: While most quantitative characters are defined in terms of maximum values, this character is defined in terms of minimum values because most songs with dissonant intervals also contain consonant intervals, but not the reverse. We do not include the standard distinctions between "consonant", "perfect" and "tritone" intervals found in Western music theory because it is not yet established to what degree these categories are cross-culturally or experimentally valid. The validity of the consonant/dissonant distinction as a function of critical bandwidth, on the other hand, has been firmly established by Plomp and Levelt (1965). Calculating critical bandwidth is impractical to do by ear, but it neatly matches the traditional division in Western music theory between "consonant" thirds and "dissonant" seconds. These labels do not imply any aesthetic value judgment.

20) RELATIVE MOTION (Cantometrics Line 22)
Relationship of the melodic contours (see **13**) of two simultaneous parts

(a) *Hetero-contour (drone):* One part is horizontal, the other changes direction (formerly "*drone polyphony*")
(b) *Poly-contour (independent motion):* Both parts have different, non-horizontal contours (formerly divided into "*harmony*" and "*counterpoint*")
(c) *Iso-contour (parallel motion):* Both parts have the same contour (formerly divided into "*isolated chords*" and "*parallel chords*")
(n/a) *One-part: S*ee **(17)**

N.B. Songs with different types of relative motion between different vocal parts or different phrases should be multi-coded.

Comments: Distinctions between different "*poly-contour*" and "*iso-contour*" sub-types (including ostinato) were removed due to their vague definitions.

III) "FORM" (between-phrase)

21) PHRASE REPETITION (Cantometrics Line 16)
Maximum number of successive phrases before a phrase is repeated

(i) *Non-repetitive*: >8 phrases, or no repeat at all (formerly "*through-composed*")
(ii) *Moderately repetitive:* 3-8 phrases (formerly "*strophe*")
(iii) *Repetitive:* 1-2 phrases (formerly "*litany")*

N.B. See Box 1 for the definition of a "phrase". Phrases where everything but the text is repeated are counted as a repeat for this character.

Comments: Because of the way phrase repetition is operationalized, the character-states are listed in an order where the number of phrases decreases rather than increases, just as they were in Cantometrics. This character no longer distinguishes between simple and complex sub-types, or between different degrees of variation in repeated phrases. (following Busby 2006). "*Canonic/round form*" and other overlapping relationships between parts are now coded in **(26)**.

22) PHRASE LENGTH (Cantometrics Line 17)
Maximum phrase length, in seconds

(i) *Short phrases:* <5 s (formerly divided into "*very short*" and "*short*" phrases)
(ii) *Medium-length phrases:* 5-9 s
(iii) *Long phrases*: >9 s (formerly divided into "*long*" and "*very long*" phrases)

Comments: As stated previously, ambiguities about where a phrase ends should be resolved by relying on breathing points to define phrase boundaries.

23) PHRASE SYMMETRY (Cantometrics Line 18)
Ratio of the length of the longest phrase in a song relative to the shortest phrase

(i) *Symmetric:* <1.5 times the length of the shortest phrase
(ii) *Mildly asymmetric*:* 1.5-2.5 times the length of the shortest phrase
(iii) *Very asymmetric**: >2.5 times the length of the shortest phrase

Comments: The original character did not define "symmetry". Characters in the original character regarding the number of phrases were removed because they were redundant with phrase repetition **(21)**.

24) SOLO/GROUP ARRANGEMENT (reorganization of Cantometrics Line 1)
Number of singers in each phrase

(a) *Solo:* Only solo phrases throughout (formerly divided into "*one solo singer*" and "*one solo singer after another*")
(b) *Mixed:* Individual phrases contain both group and solo sub-sections (formerly "*social unison with a dominant leader*")
(c) *Alternating:* Alternation between distinct solo and group phrases (formerly divided into "*simple alternation: leader-chorus*", "*overlapping alternation: leader-chorus*", and "*overlapping alternation: chorus-leader*")
(d) *Group:* Only group phrases throughout (formerly divided into "*social unison with the group dominant*", "*discoordinated*", "*simple alternation: chorus-chorus*", "*overlapping alternation: chorus-chorus*", and "*interlock*")

Comments: Cantometrics Line 1 originally contained 13 character-states that represented various complex combinations of multiple characters. Characters involving solo/group arrangement, responsorial arrangement, and phrase overlap have been moved to characters **(24)**, **(25)**, and **(26)**, respectively, to isolate the common features in these underlying characters (following Busby 2006).

25) RESPONSORIAL ARRANGEMENT (reorganization of Cantometrics Line 1)

Alternation of phrases between different vocal parts

(a) *A-responsorial:* No alternation between parts (formerly divided into "*one solo singer*", "*social unison with the group dominant*", "*discoordinated*", and "*social unison with a dominant leader*")
(b) *Hetero-responsorial*:* Irregular alternation between parts
(c) *Iso-responsorial:* Consistent alternation between parts (formerly divided into "*simple alternation: chorus-chorus*", "*overlapping alternation: chorus-chorus*", "*simple alternation: leader-chorus*", "*overlapping alternation: chorus-leader*", "*one solo singer after another*", and "*interlock*")

N.B. Songs classified as (a) ("*a-responsorial*") must be coded (n/a) for character **(26)**.

Comments: See comments in **(24)**.

26) PHRASE OVERLAP (reorganization of Cantometrics Line 1)

Maximum overlap between a "call" phrase and the "response" phrase that alternates with it (as the percentage of time in which the latter phrase overlaps with the former)

(i) *Non-overlapping:* 0% (formerly divided into "*simple alternation: chorus-chorus*", "*simple alternation: leader-chorus*", and "*one solo singer after another*")
(ii) *Mildly overlapping:* 1–25% (formerly divided into "*overlapping alternation: chorus-chorus*" and "*overlapping alternation: chorus-leader*")
(iii) *Highly overlapping:* >25% (formerly classified as "*interlock*" and/or "*canonic or round form*" in Line 16)
(n/a) *A-responsorial:* See **(25)**

Comments: See comments in **(24)**.

Sample classification

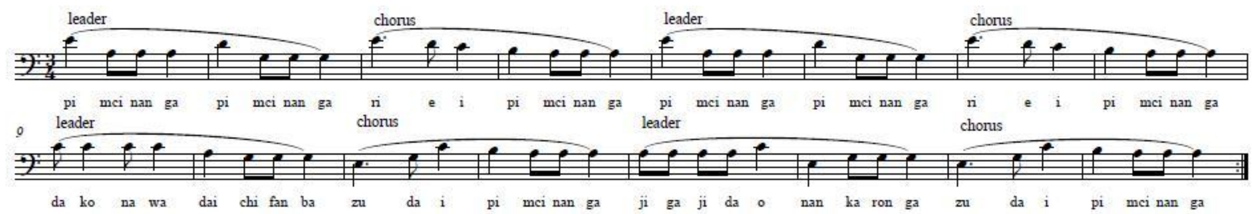
To aid in understanding the practicalities involved in applying these idealized definitions to real songs, a sample transcription of the Shona song "*Pi mcinanga*" (track 13 from Lomax's [1976] Cantometrics Consensus Tape) is provided along with a table showing how it would be classified (Fig. 4). CantoCore classifications for all 30 songs on the Cantometrics Consensus Tape are listed in Appendix A.

Character	Quantitative value	Classification
1) Meter	n/a	d) Iso-metric
2) No. of beats	n/a	b) Triple
3) Beat sub-division	n/a	c) Iso-divisive
4) No. of sub-beats	n/a	a) Simple
5) Syncopation	0%	i) Un-syncopated
6) Motivic redundancy	66% (40/61 notes derived from ♩ ♪♪♩)	iii) Highly motivic
7) Durational variability	3 unique duration values (♪, ♩, and ♩.)	ii) Moderate durational variability
8) Tonality	n/a	e) Iso-tonal
9) Mode	n/a	d) Minor iso-modal
10) No. of scale degrees	6 scale degrees (A,B,C,D,E ,G)	iii) Dense scale
11) Hemitonicity	7% (4/60 intervals are the semitone between C and B)	ii) Moderately hemitonic
12) Melodic interval size	800 cents max [C-E in bars 9-10]	iii) Large intervals
13) Melodic range	1200 cents [E-E]	ii) Medium range
14) Melodic contour	n/a	cef) Descending [phrases 2, 4 & 5]; arched [phrases 6 & 8]; undulating [phrases 1,3 & 7]
15) Melisma	1 note max	i) Syllabic
16) Vocables	11% [7/61 syllables]	i) Few vocables
17) No. of vocal parts	1	i) One-part
18) Rhythmic texture	n/a	n/a) One-part
19) Harmonic texture	n/a	n/a) One-part
20) Relative motion	n/a	n/a) One-part
21) Phrase repetition	4 phrases max before a repeat	ii) Moderately repetitive
22) Phrase length	2 seconds	i) Short phrases
23) Phrase symmetry	1 (1:1 ratio of longest:shortest phrase)	i) Symmetric
24) Solo/group arrangement	n/a	c) Alternating
25) Responsorial arrangement	n/a	c) Iso-responsorial
26) Phrase overlap	0%	i) Non-overlapping

Fig. 4: Transcription of the Shona song "*Pi mcinanga*" (track 13 from the Cantometrics Consensus Tape; mp3 available at http://greenstone.ilam.ru.ac.za/collect/ilam/index/assoc/D11847.dir/TR174-09.mp3) and its codings on the 26 CantoCore characters. The actual pitches are 2 semitones lower than those shown in the transcription. For quantitative characters, both raw quantitative values and categorical classifications are shown.

RELIABILITY

To compare the reliability of Cantometrics and CantoCore, E.M. used both systems to classify the 30 songs from the Cantometrics Consensus Tape (Fig. 5) by ear after being trained in both systems with the aid of the Cantometrics Training Tapes (Lomax 1976), but before being informed of our hypotheses about reliability. We then compared her Cantometric codings with those of the creators of Cantometrics (Lomax 1976:168-70) and her CantoCore codings with those of one of its creators (P.E.S.; his codings are shown in Appendix A). We calculated the agreement on each character separately (see Appendix B for detailed results), and then averaged across all characters to compare the mean agreement between the two classification systems.

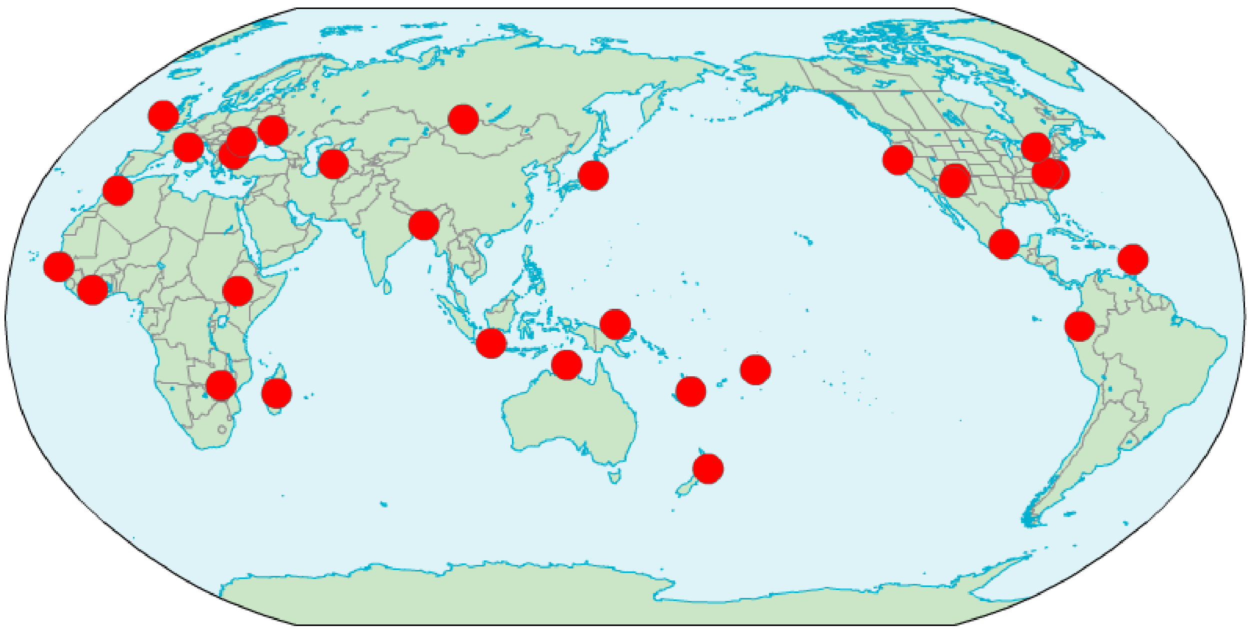


Fig. 5: Approximate geographic locations for the 30 songs from the Cantometrics Consensus Tape (Lomax 1976) used to test the reliability of Cantometrics and CantoCore. The map was generated using the *World Atlas of Language Structures Online* (http://wals.info).

We calculated inter-rater reliability for each individual character in two ways. First, we used the simplest measure, that of percent agreement. However, this statistic does not account for the effects of chance agreement, partial agreement, and character redundancy. Therefore, we also calculated reliability a second way, this time correcting for these problems using the kappa-statistic (κ), after removing all redundant codings (i.e., all "n/a" codings in CantoCore and character-state "1" ["absence"] for Cantometrics lines 2, 4-9, 12-14, 22, and 27). We used "weighted κ" (Cohen 1968) for quantitative characters and "unweighted κ" (Cohen 1960) for qualitative characters. For both percent agreement and κ, we used only the single coding indicated as most prominent in cases of multi-coding. The results for κ are shown in Figure 6.

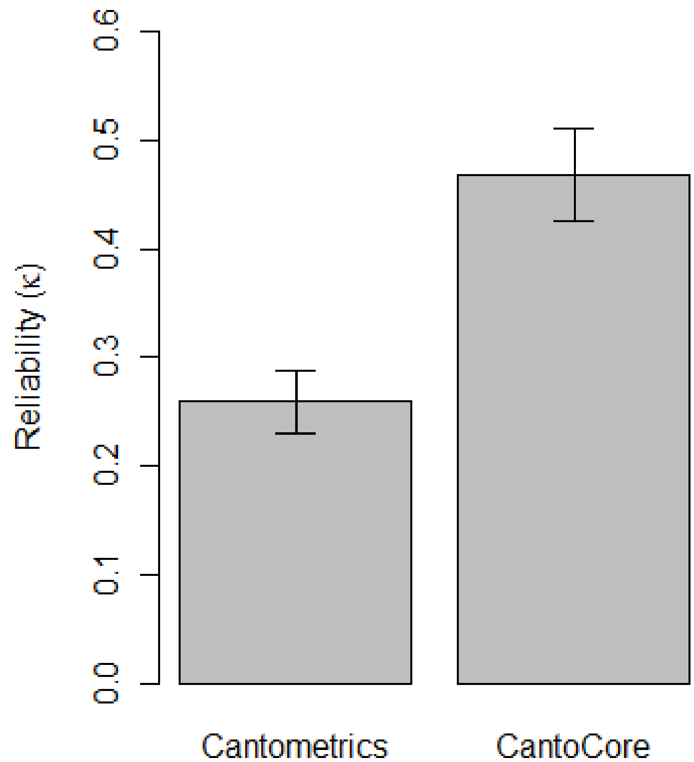


Fig. 6: Mean reliability for all 37 classification characters in Cantometrics and all 26 characters in CantoCore. Error bars represent the standard error of the mean. CantoCore is significantly more reliable than Cantometrics ($p=0.0001$).

As predicted, CantoCore appeared to be more reliable than Cantometrics. The mean percent agreement was 62% for CantoCore and 45% for Cantometrics. The results using the more precise κ statistic were highly significant statistically ($p=0.0001$), with the mean κ value of CantoCore (0.47) being approximately 80% higher than that of Cantometrics (0.26). According to Landis and Koch's (1977) criteria for interpreting κ, this translates to "moderate" reliability for CantoCore and "fair" reliability for Cantometrics, on a scale of "poor" (<0), "slight" (0-0.2), "fair" (0.21-0.4), "moderate" (0.41-0.6), "substantial" (0.61-0.8), and "almost perfect" (0.81-1). Both systems were significantly more reliable than chance ($p<1 \times 10^{-11}$), countering claims that Cantometrics is unreliable (Downey 1970; Maranda 1970; Nettl 1970).

Contrary to our predictions, there was no significant difference in reliability between the structural and performance characters of Cantometrics (structure: mean κ = 0.30, performance: mean κ = 0.29; $p=0.81$). Therefore, it may still be useful to supplement CantoCore's structural characters with the performance characters from Cantometrics. Cantometrics' instrumentation characters, however, were not significantly more reliable than chance (mean κ = 0.11, $p=0.07$), consistent with our prediction that songs are more amenable than instrumental music to reliable cross-cultural classification.

All of the reliability values for both CantoCore and Cantometrics are substantially lower than the ones given by Lomax (1976:270) and by Lomax, Halifax and Markel

(1968). However, it is difficult to compare these datasets with our results, as they used different statistics, did not present complete data or methods, and did not use a consistent song sample. At the same time, our own data should be treated as provisional, as logistical constraints limited us to collecting reliability data from only a single coder. Victor Grauer (personal communication) has pointed out that our results may reflect more on our coder and/or our training methods than on the classification schemes themselves. We accept this possibility, but maintain that we have tried our best not to bias the test in favor of CantoCore. We therefore predict that the relative reliability values of the two schemes will probably remain similar even if the absolute reliability values for both schemes is higher or lower overall for different coders. Of course, as with all science, our claims should be tested by independent researchers with larger samples to see whether they are replicable and whether they can generalize to other situations and other cultures.

APPLICATIONS

Classification is a method of examining patterns of similarity and difference. It is a means, not an end. Thus, the true test of CantoCore will be whether it, like Cantometrics, can be used as a tool to explore relationships between songs, and between music and culture. Comparing the relative similarities and differences across all CantoCore classifications can allow us to quantify different degrees and types of musical similarity. This can be used to create global musical taxonomies, in the same way that Cantometrics permitted Lomax (1968) to propose 10 canonical singing styles throughout world cultures.

The growth of the digital humanities has seen the birth of a new field of Music Information Retrieval (MIR) eager to take up the challenge of classifying music. While MIR has made great strides in adapting computational models to Western music, the lack of a theoretical framework for cross-cultural musical classification still hampers the development of “computational ethnomusicology” (Tzanetakis et al. 2007). Both CantoCore and Cantometrics provide such a framework, which computational ethnomusicologists can build on to design automated algorithms to allow for faster and more objective classification and acoustic feature-extraction. Our classification of the 30 songs from the Cantometrics Consensus Tape can act as a “ground-truth” dataset for such attempts or for atheoretical classification approaches using statistically-based machine learning algorithms. Furthermore, our improved statistical techniques for examining song similarity (Rzeszutek, Savage and Brown submitted) provide new methods for computational ethnomusicologists to analyze similarity not only between individual songs but also between diverse repertoires of heterogeneous songs.

Classification is also a tool that can be used to provide insight into musical evolution and human history. While much of the study of musical evolution has focused on music’s role in biological evolution (*why* did music evolve?) (Spencer 1857; Darwin 1871; Pinker 1997; Wallin, Merker, and Brown 2000; Cross 2001), little attention has been given to the cultural evolution of music itself (*how* does music evolve?), including the forces of musical change and stasis both geographically and historically. When attempts have been made (Lomax 1968; Lomax and Berkowitz 1972; Grauer 2006; Jan

2007), critics have rightly pointed out difficulties in distinguishing “deep” (phylogenetic) evolutionary relationships from “surface” (phenetic) acoustic similarities (Blacking 1977; Stock 2006). However, similar issues also confront the study of biological evolution (Hennig 1965; Sneath and Sokal 1973; Doolittle 1999) and cultural evolution (Mace and Holden 2005; Currie, Greenhill, and Mace 2010). Importantly, classification tools like Cantometrics and CantoCore provide a *typological* view of music – breaking music down into the principal characters that make up these schemes – and this may be useful in understanding the evolution of individual musical characters as well as elucidating musical universals (Brown and Jordania in press).

The time has come to return to Adler’s (1885) original vision of a musicology that sees classification, comparison, and ethnography as equal partners in the quest to understand the world of music. This will require using all the tools that are available, musical and non-musical, humanistic and scientific, qualitative and quantitative, theoretical and empirical. It will also require collaborative approaches that integrate work ranging from “thick description” (Geertz 1973) of individual songs or societies to “mass comparison” (Greenberg 1957) of worldwide patterns of diversity. Anthropologists have historically been split between those in the humanities who emphasize the former and those in the sciences who emphasize the latter, but there has recently been a movement towards integrating both approaches (Kuper and Marks 2011; Smith, Gurven, and Mulder 2011; Nekaris, Nijman, and Godfrey 2011). CantoCore provides a reliable method to assist in this multidisciplinary goal.

ACKNOWLEDGMENTS

We are grateful to Geoffrey Clarfield, Tom Currie, Victor Grauer, Joseph Jordania, David Locke, Michael Tenzer, and Anna Lomax Wood for critical reading of a previous version of the manuscript. We thank Sawa Matsueda Savage for the art work in Figure 1. This work was supported by the Social Sciences and Humanities Research Council (SSRHC) of Canada to S.B and by an Amherst College Roland Wood Fellowship to P.E.S.

REFERENCES

Adler, Guido. 1885. “The Scope, Method and Aim of Musicology [in German].” *Vierteljahrsschrift fur Musikwissenschaft* 1: 5-20.

Agawu, Kofi. 2010. “Against Ethnotheory.” In *Proceedings of the First International Conference on Analytic Approaches to World Music*, ed. Lawrence Shuster and Rob Schutz, 12. Amherst: Universtiy of Massachusetts.

Anku, Willie. 2000. “Circles and Time: A Theory of Structural Organization of Rhythm in African Music.” *Music Theory Online* 6.1. ISBN:1067-3040.

Atkinson, Quentin D. 2011. "Phonemic Diversity Supports a Serial Founder Effect Model of Language Expansion from Africa." *Science* 332: 346-349.

Blacking, John. 1973. *How musical is man?* Seattle: University of Washington Press.

______. 1977. "Some Problems of Theory and Method in the Study of Musical Change." *Yearbook of the International Folk Music Council* 9: 1–26.

Brown, Steven. 2000. "The 'Musilanguage' Model of Musical Evolution." In *The Origins of Music*, ed. Nils L. Wallin, Bjorn Merker, and Steven Brown, 271-300. Cambridge: MIT Press.

Brown, Steven, and Joseph Jordania. In press. "Universals in the World's Music". *Psychology of Music*.

Busby, George. 2006. "Finding the blues: An Investigation into the Origins and Evolution of African-American music." M.Sc. thesis: University of London.

Cohen, Jacob. 1960. "A Coefficient of Agreement for Nominal Scales." *Educational and Psychological Measurement* 20.1: 37-46.

______. 1968. "Weighted kappa: Nominal Scale Agreement with Provision for Scaled Disagreement or Partial Credit." *Psychological Bulletin* 70.4: 213-220.

Cross, Ian. 2001. "Music, Cognition, Culture, and Evolution." *Annals of the New York Academy of Sciences* 930: 28-42.

Currie, Thomas E., Simon J. Greenhill, and Ruth Mace. 2010. "Is Horizontal Transmission Really a Problem for Phylogenetic Comparative Methods? A Simulation Study Using Continuous Cultural Traits." *Philosophical Transactions of the Royal Society B: Biological Sciences* 365.1559: 3903-3912.

Currie, Thomas E., and Ruth Mace. 2009. "Political Complexity Predicts the Spread of Ethnolinguistic Groups." *Proceedings of the National Academy of Sciences of the United States of America* 106.18: 7339-7344.

Darwin, Charles. 1859. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray.

______. 1871. *The Descent of Man, and Selection in Relation to Sex*. London: John Murray.

Doolittle, W. Ford. 1999. "Phylogenetic Classification and the Universal Tree." *Science* 284.5423: 2124-2128.

Dowling, W. Jay, and Dane L. Harwood. 1986. *Music Cognition*. Orlando: Academic Press.

Downey, James C. 1970. "Review of A. Lomax, Folk Song Style and Culture." *Ethnomusicology* 14.1: 63-67.

Driver, Harold E. 1970. "Review of A. Lomax, Folk Song Style and Culture." *Ethnomusicology* 14.1: 57-62.

Dunn, Michael, Simon J. Greenhill, Stephen C. Levinson, and Russell D. Gray. 2011. "Evolved Structure of Language Shows Lineage-Specific Trends in Word-Order Universals." *Nature* 473: 79-82.

Ellis, Alexander J. 1885. "On the Musical Scales of Various Nations." *Journal of the Society of Arts* 33.1: 485-527.

Erickson, Edwin E. 1976. "Tradition and Evolution in Song Style: A Reanalysis of Cantometric Data." *Cross-Cultural Research* 11.4: 277-308.

Feld, Steven, and Aaron A. Fox. 1994. "Music and Language." *Annual Review of Anthropology* 23: 25-53.

Geertz, Clifford. 1973. *The Interpretation of Cultures*. New York: Basic Books.

Grauer, Victor. 2005. "'Cantometrics – Song and Social Culture': A Response." *Musical Traditions* 159. http://www.mustrad.org.uk/articles/cantome2.htm.

Grauer, Victor. 2006. "Echoes of our Forgotten Ancestors." *The World of Music* 48.2: 5-59.

Greenberg, Joseph H. 1957. *Essays in Linguistics*. Chicago: University of Chicago Press.

Harris, Marvin. 1976. "History and Significance of the Emic/Etic Distinction." *Annual Review of Anthropology* 5.1: 329-350.

Haspelmath, Martin, Matthew S. Dryer, David Gil, and Bernard Comrie, eds. 2005. *The World Atlas of Language Structures*. Oxford: Oxford University Press.

Hennig, Willi. 1965. "Phylogenetic Systematics." *Annual Review of Entomology* 10.1: 97–116.

Henry, Edward O. 1976. "The Variety of Music in a North Indian Village: Reassessing Cantometrics." *Ethnomusicology* 20.1: 49-66.

Hood, Mantle. 1971. *The Ethnomusicologist*. New York: McGraw-Hill.

Hornbostel, Erich M. von, and Curt Sachs. 1914. “Classification of Musical Instruments [in German].” *Zeitschrift fur Ethnologie* 14: 3-29.

Huron, David. 2006. *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge: MIT Press.

Jan, Steven. 2007. *The Memetics of Music: A Neo-Darwinian View of Musical Structure and Culture*. Hants: Ashgate.

Jones, William. 1807. *The Works of Sir William Jones*. 1st ed. London: John Hatchard.

Kolinski, Mieczyslaw. 1961. “Classification of Tonal Structures.” *Studies in Ethnomusicology* 1: 38-76.

______. 1962. “Consonance and Dissonance.” *Ethnomusicology* 6.2: 66–74.

______. 1973. “A Cross-Cultural Approach to Metro-Rhythmic Patterns.” *Ethnomusicology* 17.3: 494-506.

Krumhansl, Carol. 1990. *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press.

Kuper, Adam, and Jonathan Marks. 2011. “Anthropologists Unite!” *Nature* 470.7333: 166-168.

Landis, J. Richard, and Gary G. Koch. 1977. “The Measurement of Observer Agreement for Categorical Data.” *Biometrics* 33.1: 159-174.

Lerdahl, Fred, and Ray Jackendoff. 1983. *A Generative Theory of Tonal Music*. Cambridge: MIT Press.

Leroi, Armand M., and Jonathan Swire. 2006. “The Recovery of the Past.” *The World of Music* 48.3: 43-54.

Lewis, M. Paul, ed. 2009. *Ethnologue: Languages of the World*. 16th ed. Dallas: SIL International.

Lomax, Alan. 1959. “Folk Song Style.” *American Anthropologist* 61.6: 927–954.

______. 1976. *Cantometrics: An Approach to the Anthropology of Music*. Berkeley: University of California Extension Media Center.

Lomax, Alan, and Norman Berkowitz. 1972. “The Evolutionary Taxonomy of Culture.” *Science* 177.4045: 228-239.

Lomax, Alan, ed. 1968. *Folk Song Style and Culture*. Washington, DC: American Association for the Advancement of Science.

Lomax, Alan, and Victor Grauer. 1968. "The Cantometric Coding Book." In *Folk Song Style and Culture*, ed. Alan Lomax, 34-74. Washington, DC: American Association for the Advancement of Science.

Lomax, Alan, Joan Halifax, and Norman N. Markel. 1968. "Consensus on Cantometric Parameters." In *Folk Song Style and Culture*, ed. Alan Lomax, 111-116. Washington, DC: American Association for the Advancement of Science.

Mace, Ruth, and Clare J. Holden. 2005. "A Phylogenetic Approach to Cultural Evolution." *Trends in Ecology and Evolution* 20.3: 116-21.

Maranda, Elli Kongas. 1970. "Deep Significance and Surface Significance: Is Cantometrics Possible?" *Semiotica* 2.2: 173-184.

McLeod, Norma. 1974. "Ethnomusicological Research and Anthropology." *Annual Review of Anthropology* 3.1: 99-115.

Merriam, Alan P. 1982. "On Objections to Comparison in Ethnomusicology." In *Cross-cultural perspectives on music*, ed. Robert Falck and Timothy Rice, 175-189. Toronto: University of Toronto Press.

Naroll, Raoul. 1969. "Singing and Dancing: A Cross-Cultural Survey." *Science* 166.3903: 366-367.

Nekaris, K. Anne-Isola, Vincent Nijman, and Laurie R. Godfrey. 2011. "Anthropology: Follow Field Primatologists." *Nature* 471: 448.

Nettl, Bruno. 1970. "Review of A. Lomax, Folk Song Style and Culture." *American Anthropologist* 72.2: 438-441.

______. 2005. *The Study of Ethnomusicology: Thirty-One Issues and Concepts*. 2nd ed. Champaign: University of Illinois Press.

Nettl, Bruno, Ruth Stone, James Porter, and Tim Rice. 1998. *The Garland Encyclopedia of World Music (10 vols.)*. New York: Routledge.

Patel, Aniruddh D. 2008. *Music, Language and the Brain*. Oxford: Oxford University Press.

Patel, Aniruddh D., and Joseph R. Daniele. 2003. "An Empirical Comparison of Rhythm in Language and Music." *Cognition* 87.1: 35-45.

Pfordresher, Peter Q., Steven Brown, Kimberly M. Meier, Michel Belyk, and Mario Liotti. 2010. "Imprecise Singing is Widespread." *The Journal of the Acoustical Society of America* 128.4: 2182-2190.

Pinker, Steven. 1997. *How the Mind Works*. New York: Norton.

Plomp, R., and W.J.M. Levelt. 1965. "Tonal Consonance and Critical Bandwidth." *The Journal of the Acoustical Society of America* 38: 548-560.

Rzeszutek, Tom, Patrick E. Savage, and Steven Brown. Submitted. "The Structure of Cross-Cultural Musical Diversity." *Proceedings of the Royal Society B: Biological Sciences*

Sachs, Curt. 1943. *The Rise of Music in the Ancient World: East and West*. New York: Norton.

Schenker, Heinrich. 1979. *Free Composition (Der freie Satz): Volume III of New Musical Theories and Fantasies*. New York: Longman.

Smith, Eric Alden, Michael Gurven, and Monique Borgerhoff Mulder. 2011. "Anthropology: It Can be Interdisciplinary." *Nature* 471: 448.

Sneath, Peter H.A., and Robert R. Sokal. 1973. *Numerical Taxonomy: The Principles and Practice of Numerical Classification*. San Francisco: W.H. Freeman and Company.

Spencer, Herbert. 1857. "The Origin and Function of Music." *Fraser's Magazine* 56: 396-408.

Stock, J P J. 2006. "Clues from our Present Peers? A Response to Victor Grauer." *The World of Music* 48.2: 73-91.

Tenzer, Michael, ed. 2006. *Analytical Studies in World Music*. New York: Oxford University Press.

Toiviainen, Petri, and Tuomas Eerola. 2001. "A Method for Comparative Analysis of Folk Music Based on Musical Feature Extraction and Neural Networks." In *Proceedings of the VII International Symposium of Systematic and Comparative Musicology and the III International Conference on Cognitive Musicology*, ed. H. Lappalainen, 41-45. Jyväskylä: University of Jyväskylä.

Tzanetakis, George, Ajay Kapur, W. Andrew Schloss, and Matthew Wright. 2007. "Computational Ethnomusicology." *Journal of Interdisciplinary Music Studies* 1.2: 1-24.

Wallin, Nils L., Bjorn Merker, and Steven Brown, eds. 2000. *The Origins of Music*. Cambridge: MIT Press.

APPENDICES

Appendix A: Sample CantoCore codings

Table A1. CantoCore codings of all 30 songs from the Cantometrics Consensus Tape, as done by P.E.S. The 30 songs (see Lomax 1976:164-171 for details) are listed by row number, and the 26 CantoCore characters are listed by column number. In cases of multi-coding, the most prominent coding is bolded. See text for a detailed description of the characters and character-states.

Cantometrics Consensus Tape track number	**CantoCore character number**																									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**
1	a**d**	a	c	a	ii	iii	iii	e	c**e**	ii	ii	ii	iii	ac**e**f	i	ii	iii	**b**c	ii	a**b**c	ii	ii	iii	b	b	i
2	d	a	c	a	i	iii	ii	e	e	ii	i	ii	iii	f	i	ii	i	n/a	n/a	n/a	iii	ii	i	a	a	n/a
3	d	a	c	a	ii	iii	i	e	e	ii	ii	ii	ii	a**c**	i	iii	i	n/a	n/a	n/a	ii	ii	i	d	a	n/a
4	a	n/a	n/a	n/a	n/a	ii	iii	e	e	iii	ii	ii	iii	f	iii	i	i	n/a	n/a	n/a	i	ii	i	a	a	n/a
5	d	a	c	a	ii	iii	ii	e	e	ii	i	iii	iii	**c**f	i	i	i	n/a	n/a	n/a	ii	i	i	a	a	n/a
6	d	a	c	a	iii	iii	ii	e	e	ii	iii	ii	iii	e	iii	i	i	n/a	n/a	n/a	ii	ii	ii	a	a	n/a
7	a	n/a	n/a	n/a	n/a	i	iii	e	e	ii	ii	ii	i	a**f**	iii	i	ii	b**c**	i	a	ii	iii	iii	d	a	n/a
8	b	n/a	n/a	n/a	i	ii	ii	e	e	ii	i	iii	ii	**a**e	iii	iii	i	n/a	n/a	n/a	ii	ii	ii	d	a	n/a
9	d	a	c	b	i	iii	ii	e	e	iii	ii	iii	iii	**e**f	i	i	i	n/a	n/a	n/a	iii	ii	i	a	c	ii
10	d	c	b	n/a	ii	iii	ii	e	d	iii	ii	ii	iii	**c**e	i	i	i	n/a	n/a	n/a	ii	i	i	a	a	n/a
11	b	n/a	n/a	n/a	i	ii	i	e	d	ii	ii	ii	ii	f	i	iii	ii	c	ii	c	iii	ii	i	b	b	i
12	b	n/a	n/a	n/a	iii	ii	i	a**e**	d	ii	i	ii	iii	**a**c	ii	iii	ii	b	n/a	a	ii	ii	iii	b	a	n/a
13	d	b	c	a	i	iii	ii	e	d	iii	ii	iii	ii	**c**ef	i	i	i	n/a	n/a	n/a	ii	i	i	c	c	i
14	d	a	c	a	i	iii	ii	a**e**	d	ii	i	ii	iii	**e**f	ii	i	i	n/a	n/a	n/a	ii	ii	ii	a	a	n/a
15	d	a	c	a	ii	iii	ii	e	e	iii	ii	ii	iii	c	i	i	ii	c	ii	c	iii	ii	i	c	c	i
16	b	n/a	n/a	n/a	i	ii	ii	e	d	ii	ii	ii	iii	c	i	ii	i	n/a	n/a	n/a	iii	ii	ii	d	a	n/a
17	d	a	c	a	i	iii	ii	e	e	i	i	iii	iii	f	i	i	i	n/a	n/a	n/a	iii	i	i	a	a	n/a
18	a	n/a	n/a	n/a	n/a	ii	ii	e	a	ii	iii	i	i	a	iii	i	i	n/a	n/a	n/a	iii	ii	i	d	a	n/a
19	a	n/a	n/a	n/a	n/a	iii	ii	e	d	i	ii	ii	i	e	ii	iii	i	n/a	n/a	n/a	iii	i	i	c	c	i
20	d	a	c	b	i	iii	ii	e	e	ii	i	ii	iii	f	iii	iii	ii	b	n/a	a	iii	ii	i	a	a	n/a
21	b	n/a	n/a	n/a	i	ii	ii	c	a	ii	iii	i	i	**a**c	iii	ii	i	n/a	n/a	n/a	i	ii	ii	a	a	n/a
22	a	n/a	n/a	n/a	n/a	ii	iii	e	d	ii	ii	ii	iii	a**f**	iii	i	i	n/a	n/a	n/a	i	ii	ii	a	a	n/a
23	d	b	c	a	ii	iii	ii	e	e	iii	ii	ii	ii	de**f**	i	i	iii	c	ii	a**b**c	ii	ii	i	d	a	n/a
24	d	a	c	a	ii	iii	ii	e	e	ii	i	i	i	a	i	ii	ii	c	i	**a**c	iii	i	i	d	a	n/a
25	b	n/a	n/a	n/a	i	ii	i	e	d	iii	ii	i	ii	c	i	iii	ii	b	n/a	a	iii	ii	i	b	b	iii
26	b	n/a	n/a	n/a	ii	iii	ii	e	d	i	ii	ii	ii	a**c**	iii	iii	i	n/a	n/a	n/a	iii	ii	iii	d	a	n/a
27	d	a	c	b	ii	iii	ii	e	a	i	i	ii	iii	a	i	iii	iii	b	n/a	b	iii	i	i	d	c	iii
28	d	c	b	n/a	i	ii	ii	e	e	ii	i	ii	ii	c**f**	i	ii	i	n/a	n/a	n/a	iii	ii	i	d	c	i
29	c	n/a	n/a	n/a	i	iii	iii	e	a	ii	i	i	i	a	iii	iii	iii	c	ii	c	iii	i	i	d	c	ii
30	b	n/a	n/a	n/a	ii	iii	iii	e	e	iii	ii	ii	ii	**a**e	i	i	ii	c	ii	c	ii	ii	ii	d	b	ii

Appendix B: Inter-rater reliability

Table B1. Inter-rater reliability values for *song-structure* characters from CantoCore and Cantometrics. See text for a description of κ as a measurement of reliability.

Character	**Line number**		**Reliability (κ)**	
	CantoCore	**Cantometrics**	**CantoCore**	**Cantometrics**
Meter	1	11	0.36	0.043
Number of beats	2	n/a	0.60	n/a
Beat sub-division	3	n/a	0.08	n/a
Numbe of sub-beats	4	n/a	undefined	n/a
Syncopation	5	n/a	0.35	n/a
Motivic redundancy	6	n/a	0.22	n/a
Durational variability	7	n/a	0.32	n/a
Tonality	8	n/a	undefined	n/a
Mode	9	n/a	0.25	n/a
Number of scale degrees	10	n/a	0.38	n/a
Hemitonicity	11	n/a	0.20	n/a
Melodic interval size	12	21	0.48	0.36
Melodic range	13	20	0.40	0.33
Melodic contour	14	15	0.37	0.19
Melisma	15	29	0.81	0.42
Vocables	16	10	0.53	0.62
Number of vocal parts	17	4	0.68	0.15
Rhythmic texture	18	12	0.62	0.33
Harmonic texture	19	n/a	1.00	n/a
Relative motion	20	22	0.25	0.14
Phrase repetition	21	16	0.70	0.23
Phrase length	22	17	0.39	0.22
Phrase symmetry	23	18	0.51	0.35
Solo/group arrangement	24	1	0.62	0.48
Responsorial arrangement	25	n/a	0.50	n/a
Phrase overlap	26	n/a	0.64	n/a
Position of the final tone	n/a	19	n/a	0.36

Table B2. Inter-rater reliability values for *performance-style* characters (Cantometrics only). See text for a description of κ as a measurement of reliability.

Character	Line number	Reliability (κ)
Tonal blend	5	0.49
Rhythmic blend	6	0.29
Embellishment	23	0.19
Tempo	24	0.14
Volume	25	0.66
Rubato	26	0.31
Glissando	28	0.13
Tremolo	30	0.46
Glottal shake	31	0.20
Register	32	0.25
Vocal width	33	0.28
Nasalization	34	0.15
Raspiness	35	0.12
Accent	36	0.24
Enunciation	37	0.42

Table B3. Inter-rater reliability values for *instrumentation* characters (Cantometrics only). See text for a description of κ as a measurement of reliability.

Character	Line number	Reliability (κ)
Relationship to voice	2	0.14
Responsorial arrangement	3	0.22
Number of instrumental parts	7	0.07
Tonal blend	8	omitted (Lomax 1976)
Rhythmic blend	9	-0.20
Meter	13	0.34
Rhythmic texture	14	0.16
Rubato	27	0.04

Chapter 3

The structure of cross-cultural musical diversity

Tom Rzeszutek, Patrick E. Savage, and Steven Brown

Department of Psychology, Neuroscience and Behaviour, McMaster University, Hamilton, ON, Canada

Keywords: AMOVA; cultural diversity; music; population structure; Austronesian language family; Taiwan

3.1 Abstract

Human cultural traits, such as languages, musics, rituals, and material objects, vary widely across cultures. However, the majority of comparative analyses of human cultural diversity focus on between-culture variation without consideration for within-culture variation. In contrast, biological approaches to genetic diversity, such as the Analysis of Molecular Variance (AMOVA) framework, partition genetic diversity into both within- and between-population components. We attempt here for the first time to quantify both components of cultural diversity by applying the AMOVA model to music. By applying this approach to 421 traditional songs from 16 Austronesian-speaking populations, we show that the vast majority of musical variability is due to differences within populations rather than differences between. This demonstrates a striking parallel between the structures of cultural and genetic diversity in humans. A Neighbor-Net analysis of pairwise population musical divergence shows a large amount of reticulation, indicating the pervasive occurrence of borrowing and/or convergence of musical features across populations.

3.2 Introduction

Human cultural traits exhibit an astounding myriad of forms, perhaps best exemplified by the approximately 6900 known languages currently spoken across the world [1]. Any approach to characterising this cross-cultural diversity depends on the creation of a reliable classification of forms for a given domain of culture. There are many important examples of cultural classification, spanning from the seminal work of Murdock on the classification of over 100 categories of cultural behaviour across 1100 world populations [2] to contemporary examples in linguistics such as the World Atlas of

Language Structures [3] and the Austronesian Basic Vocabulary Database [4]. The primary goal of these kinds of classification systems is the identification of salient differences between populations, as these differences can aid in reconstructing the history of human population movements and cultural interactions [5-7]. A major criticism of these approaches, though, is that they place an exclusive emphasis on the diversity *between* cultures, downplaying or ignoring the internal diversity present within each culture. Overall, there is a dichotomy between comparative approaches – whose goal is to characterize differences between cultures – and ethnographic approaches, whose goal is to rigorously catalogue the richness of forms that exist within single cultures. Here, we propose a compromise solution that allows for the simultaneous consideration of between-culture and within-culture facets of cultural diversity.

The hierarchical structure of human cultural diversity is reminiscent of the structure of human genetic diversity in that diversity can be compartmentalized into within- and between-population components. Population geneticists, starting with Lewontin [8], have repeatedly observed that the vast majority of the genetic diversity in human populations is found within populations rather than between them [9]. Some cultural scholars have argued that human cultures exhibit a much lower level of internal diversity than that seen in the genetic domain due to processes such as conformity or frequency-dependent selection [10] that homogenize behaviours within populations and thereby push particular cultural variants to fixation [11]. While this is a plausible argument, no one, to the best of our knowledge, has done a rigorous quantification of the hierarchical structure of cultural diversity. Perhaps the closest is a study by Bell et al. [12] that used internal behavioural variation to calculate cultural variation among populations using a population genetic model, but this work did not explicitly quantify the degree of internal variation.

One requirement in applying population genetic models to cultural forms is the necessity that there be quantifiable features, which vary among individuals or entities both within and between populations. For example, Bell et al. [12] used questions from the World Values Survey, administered to a sample of individuals from each focal culture. This is comparable to looking at variation among individuals at a particular genetic locus. Alternatively, if one wanted to investigate variation in some aspect of material culture, such as ceramics, one would need a number of exemplars from each culture, appropriate features to describe said exemplars, and a suitable quantitative measure of differences among entities. Clearly, there is a difference between studying variation among individuals in terms of behaviour and variation among entities of material culture. What is most important for the study of cultural diversity is that the unit of analysis and the means of measuring difference between cultural variants have domain-specific validity, and this must be worked out on a case-by-case basis for each domain of culture.

Music seems to satisfy these important requirements and thereby affords a novel opportunity to study the structure of cultural diversity. Not only is music a human universal [13] but its form also varies quite prominently both between [14] and within [15] cultures. Musical features are also quite amenable to comparative analysis [14]. Most importantly for our purposes, the "song" provides a reliable unit for the cultural analysis of music. Ethnographic analyses of human cultures have clearly shown that the song represents the fundamental unit of both structure and function [13]. In addition, the song was adopted as the unit of analysis in the most ambitious comparative attempt to classify the world's musics, namely Lomax's Cantometrics project of the 1960's [14], in which more than 4000 songs from over 200 cultures were analysed and compared.

In order to make such a global project feasible, Lomax employed a small sample of only ten songs per culture, and these were averaged into a 'modal profile' that represented the 'typical' song-style for each culture [14]. While Lomax believed that his modal profiles were representative of the cultures he was sampling, ethnographers studying musics from those same cultures questioned Lomax's findings, because his approach strongly underestimated the degree of internal musical diversity in those cultures [15,16]. To date, there has been no quantitative method applied to music that retains the cross-cultural scope of Lomax's global framework while at the same time taking internal variation into account.

Exactly such a method is used in the study of genetic diversity in population genetics, and this method provides a promising approach for thinking about the hierarchical structure of cultural diversity as well. The Analysis of Molecular Variance (AMOVA) is a method closely related to the Analysis of Variance (ANOVA) that allows the hierarchical partitioning of genetic variance into components [17]. These components generally include: variability within populations; variability between populations; and variability between regional groups. The population structure being tested is defined a priori by the researcher, and can include divisions based on geographic region or language [17]. In its original application, AMOVA was designed to investigate molecular diversity based on haplotype restriction polymorphism data, but the generalizability of the method was recognized early on [17] and has since been applied to many different kinds of genetic loci [18]. The flexibility of this method rests on the fact that variability is calculated as a measure of distance between haplotypes. The distance measure itself is defined by the user and can incorporate information about sequence evolution such as mutation rate [17]. Consequently, given an appropriate unit of analysis and distance measurement, this method can be extended to quantify the hierarchical structure of cultural diversity.

We attempt here for the first time to quantify both the within- and between-population components of cultural diversity by applying AMOVA to the analysis of musical diversity using the song as the unit of analysis. An important distinction here is that we are looking at populations of *songs* rather than populations of individuals. To this end, we focus on a rigorous sampling of tribal musics from Austronesian-speaking populations in Taiwan and the Philippines, itself part of a larger project devoted to prehistoric migrations in the region. To quantify musical variability, we calculate the distance between songs using a musical-classification system we developed that is inspired by Cantometrics. The AMOVA framework is then applied to this data in order to apportion musical variability into within- and between-population components. We also measure pairwise population musical divergence with Φ_{ST} and use it in a 'Neighbor-Net' analysis [19] to explore the degree of reticulation in the data due to borrowing and/or convergence. Distances based on Φ_{ST} are also compared with the corresponding modal profiles to test the accuracy of Lomax's modal-profile approach for distinguishing differences between populations. Our novel application of AMOVA to cultural forms provides a general means of performing population-level cultural analyses while simultaneously addressing the internal diversity of cultural forms.

3.3 Materials and Methods

a) Sample

The musical sample consists of 421 traditional group (choral) songs from 16 Austronesian-speaking aboriginal populations from Taiwan and the northern Philippines, including the Amis (30 songs), Atayal (10), Bunun (30), Paiwan (30), Puyuma (30), Rukai (30), Saisiyat (30), Tao (30), Tsou (22), Plains (Siraya) (24), Kavalan (18), Thao (30), Ibaloi (30), Ifugao (30), Kankanai (17), and Ayta (30). Songs were obtained from commercial ethnomusicology recordings as well as from the Taiwan National Music Archive in Taipei [20] and the Centre for Ethnomusicology at the University of the Philippines in Quezon City. 30 songs were randomly sampled from each population. For populations with less than 30 songs available, all recordings meeting our inclusion criteria were used.

b) Classifying songs

P.E.S. coded all the songs using the "CantoCore" song-classification scheme developed in our lab (Savage, Merritt, Rzeszutek and Brown, in preparation). This comprehensive scheme, modelled after Lomax and Grauer's (1968) original Cantometric scheme [14], codes 26 characters related to song structure, including rhythm, pitch, syllable, texture, and form (see electronic supplementary material, S1).

c) Quantifying musical distance

Either phylogenetic distances based on sequence evolution or phenetic distances based on sequence similarity can be used in genetic analyses [17]. Since we currently lack information about song evolution, we attempted to develop a simple phenetic measure of distance between songs, based on our coding, that is both musically and statistically valid. Leroi and Swire [21] and Busby [22] identified a number of methodological solutions to issues related to converting Cantometric song-codings into distances, and these issues apply equally to CantoCore. These include: the presence of both ordinal and nominal characters; simultaneous coding of multiple states for a number of characters (multi-coding); the redundancy of some codings when certain states are absent; and equal weighting of all characters. We built on their work to program an algorithm that takes these issues into account while at the same time being flexible enough to handle a variety of coding schemes. The algorithm was programmed in R version 2.12.2 [23] by T.R. and is available upon request. Details of the algorithm are found in the electronic supplementary material (S2).

d) Visualizing song relationships

In order to visualize songs in two dimensions we performed nonmetric multidimensional scaling on the song-level distances obtained from our algorithm using isoMDS in R, with

50 iterations and metric scaling as an initial configuration.

e) AMOVA analysis

Distances were prepared for the AMOVA analysis by a Euclidean transform of the data using Lingoes' method [24], as implemented in the ade4 package for R [25]. The distances were then squared, as recommended by Excoffier et al. [17]. AMOVA was performed in Arlequin 3.11 using the prepared distance matrix and standard settings [26]. Musical variability was apportioned "between" and "within" ethno-linguistically defined populations of songs [1]. The parameter Φ_{ST} is the proportion of total variability due to differences between populations [17], and it was calculated pairwise as a measure of musical divergence between populations. To test the significance of the between-population component of musical variance, songs were permuted randomly between populations, using 1000 permutations.

f) Neighbor-Net analysis

Pairwise Φ_{ST}, was used in a 'Neighbor-Net' analysis [19] to determine the level of reticulation in the data due to borrowing and convergence. The analysis was performed in SplitsTree4 using standard settings [27]. All negative Φ_{ST} values were set to zero before performing the analysis [28].

g) Modal Profile Analysis

In order to test the efficacy of Lomax's modal profile approach at distinguishing differences between populations, we created a modal song coding for each population, consisting of the most common coding in its musical repertoire for each of the 26 "CantoCore" characters. This resulted in some combinations not present in any single song, but nonetheless best approximates Lomax's creation of 'modal profiles'. These modal profiles are available in the electronic supplementary material (S3). Distances between modal profiles were calculated using the same algorithm applied to the original song data, giving us a population-level distance devoid of any information about internal diversity. These modal distances were then compared to population pairwise Φ_{ST} using Spearman's rho and a Mantel test with 20000 permutations.

3.4 Results

a) Multidimensional Scaling

Figure 1 shows a multidimensional scaling plot for the 421 songs used in our sample colour-coded for the 16 tribes. The high level of stress in this two-dimensional ordination indicates the complex multidimensional nature of the musical data. A scree plot did not reveal a clear elbow, and showed that instead our data would require more than 8 dimensions to achieve an acceptable level of stress under 10. Despite this, the multidimensional scaling plot clearly demonstrates the high level of internal heterogeneity in each population's musical repertoire, and the high degree of overlap between populations.

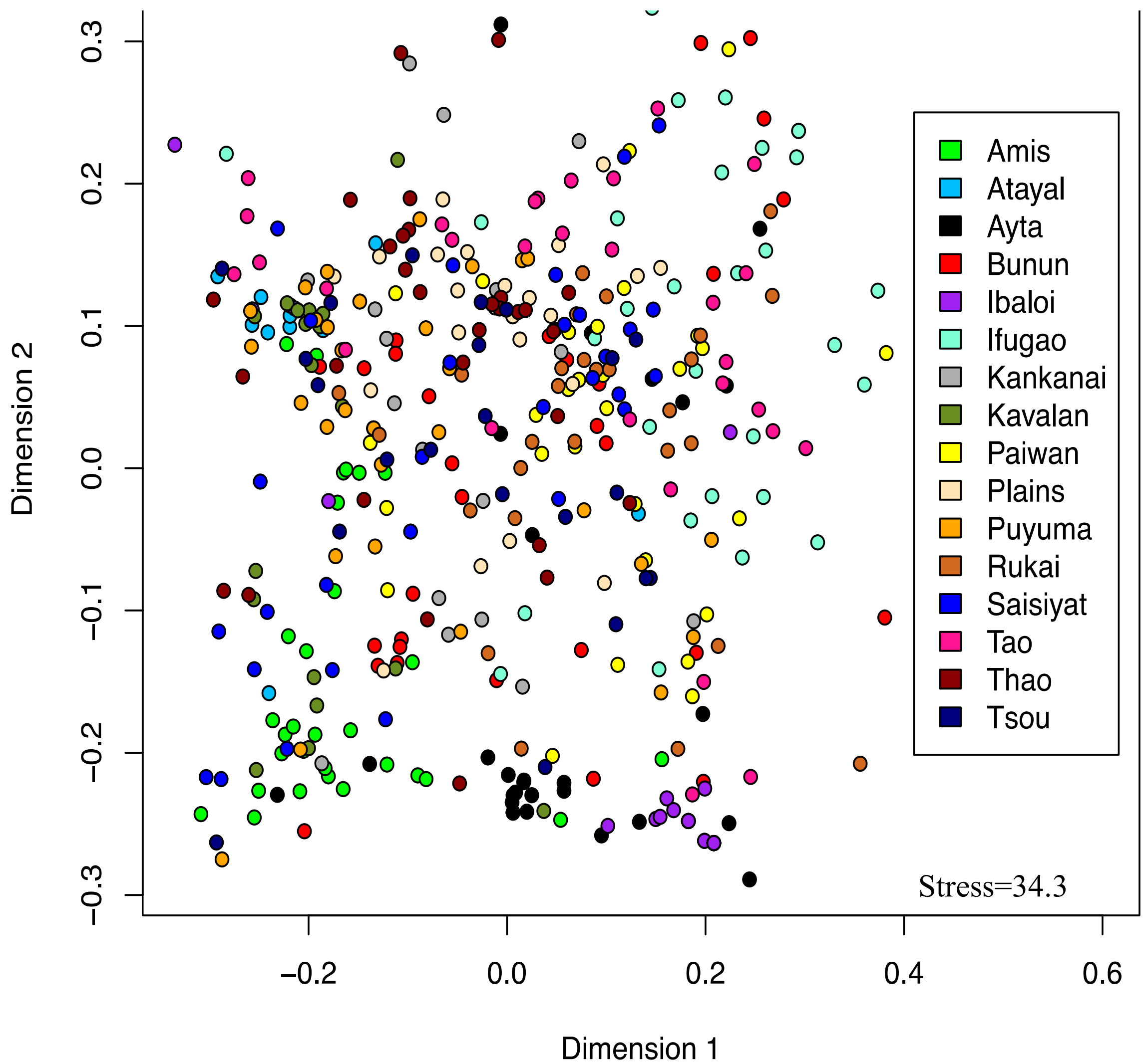


Figure 3.1 - Multidimensional scaling plot of distances between 421 songs from Austronesian-speaking populations. There is a large amount of overlap between populations and spread within populations. Each point represents a song and is coloured according to population of origin.

Table 1.
Musical AMOVA Results

Source of Variation	Percentage of Variation	Degrees of Freedom
Between Populations	2.06	15
Within Populations	97.94	405
Total	100	420

b) Song-level AMOVA analysis

The AMOVA analysis confirms the multidimensional scaling result (table 1), with a majority of the variance in our sample (~98%) being accounted for by differences within populations and a smaller portion (~2%) accounting for differences between populations. Despite accounting for a much smaller proportion of the variance, musical diversity between populations was statistically significant (Φ_{ST}=0.02057, p<0.001).

c) Neighbor-Net analysis

The Neighbor-Net analysis (figure 2) demonstrated that our musical data did not appear tree-like and instead contained a fair amount of reticulation. This reticulation is reflective of borrowing and/or convergence of musical features between populations.

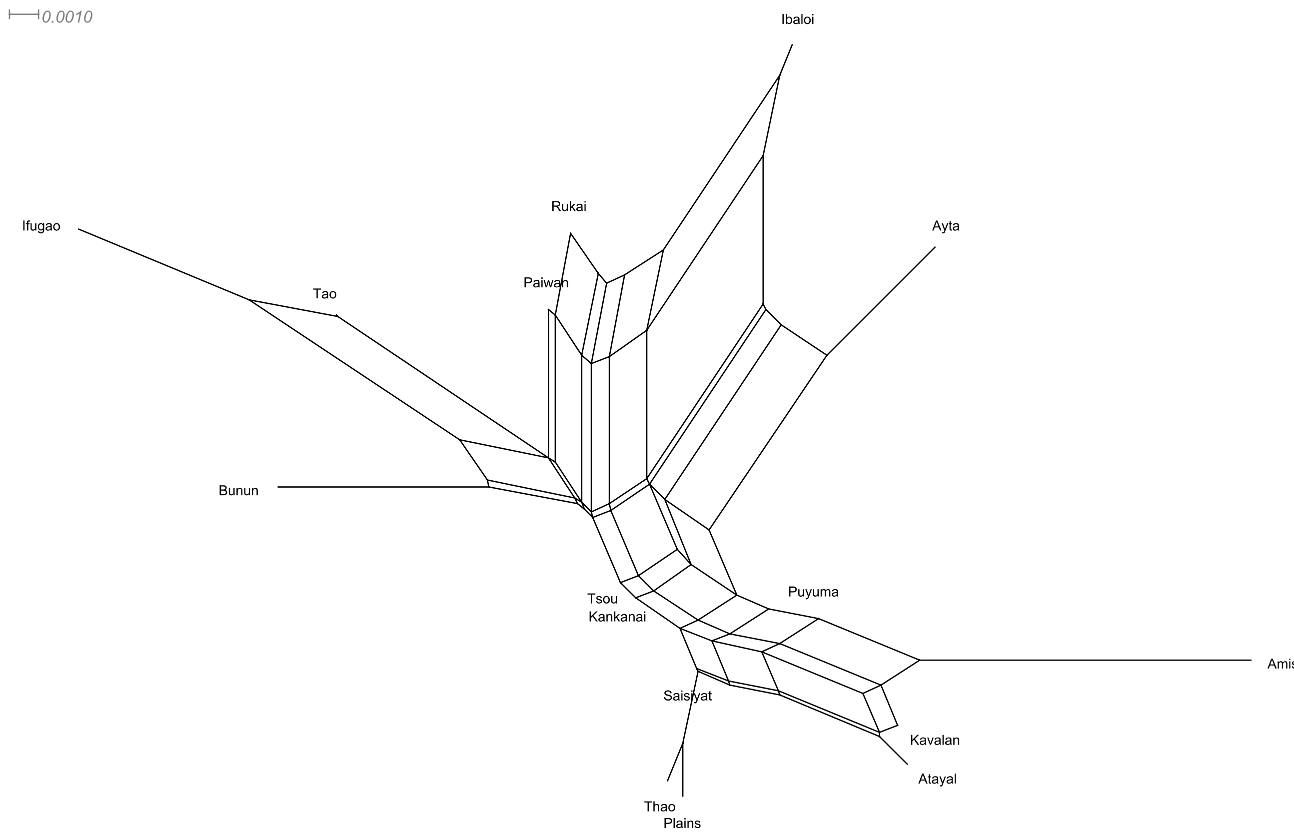


Figure 2. A Neighbor-Net plot of population-level musical divergence between 16 Austronesian-speaking populations, based on pairwise Φ_{ST} from and AMOVA analysis of 421 traditional group songs. This plot shows a high degree reticulation in our dataset, indicating the presence of borrowing and/or convergence.

d) Modal profile analysis

The pairwise population-level distances based on the modal profiles (ignoring internal diversity) were highly correlated with pairwise Φ_{ST} distances (r^2=0.533, p<0.001), which take into account the internal variation in musical repertoires. This indicates that, although it cannot capture information about internal diversity within cultures, the modal profile approach may still adequately approximate overall patterns of variation between populations (see electronic supplementary material, S4).

3.5 Discussion

We have applied the Analysis-of-Molecular-Variance (AMOVA) framework to a cultural dataset, allowing us for the first time to quantify the hierarchical structure of cultural diversity. Our application of this approach to a sample of aboriginal Austronesian songs demonstrated that the vast majority of musical variation in this sample (~98%) was found within populations, while a far smaller proportion of this variation (~2%) occurred between populations. This validates and quantifies the critiques of ethnomusicologists that Cantometrics' cross-cultural approach underestimated the diversity of musical repertoires within each culture [15,16]. Next, a Neighbor-Net analysis of population pairwise Φ_{ST} distances showed that our musical data was not very tree-like, providing some preliminary insight into the evolution of musical repertoires and the presence of forces that diversify musics within cultures.

a) How much diversity is sufficient?

The high level of internal musical diversity found in this study parallels general findings on the structure of human genetic diversity, with some estimates of this diversity being as high as 93-95% [9]. However, as in genetic domain, this raises the important question of how much diversity is sufficient for describing differences between populations. This has been extensively addressed in population genetics. Lewontin's 1972 analysis of human genetic variation led him to argue that the small proportion of variation found between populations in his study (14.6%) meant that differences between populations were not informative [8]. Some scholars [29,30], most prominently Edwards [31], have noted that this conclusion is statistically inaccurate, as it ignores information contained in the correlation of allele frequencies across many loci. Modern clustering approaches use the correlated nature of genetic data to distinguish between major human groups that coincide with their geographic distribution, despite the small amount of variation (3-5%) accounting for these differences [9].

This situation is qualitatively the same in the study of musical diversity, since the correlation between different musical features in songs reveals much more about the unique musical repertoires of populations than the frequency of the features themselves. Therefore, our observation that between-population musical variance is a very small proportion of the total variance in no way precludes using this component for taxonomic and comparative analyses of world musics, as Lomax did [14]. This kind of comparative methodology should not be applied recklessly but in consultation with expert ethnomusicologists, who can vouch for the validity of the sample. The between-population component should be sufficient to distinguish populations musically, and this is validated by our modal-profile analysis. That analysis demonstrated that a methodology that ignores internal diversity might still be successful at detecting the overall pattern of differences between populations, even though it might fail to detect more fine-grained population relationships.

b) Cultural evolution of music

The transmission of cultural traits is distinct from that of biological traits in that there are many more possible modes of transmission. Unlike in the human-genetic domain, where variants are passed vertically across generations, features of culture can also pass horizontally between members of the same cohort, as well as obliquely from unrelated elder members of a focal individual's group [32]. The presence of alternative modes of transmission has been a central issue in the application of phylogenetic models to cultural traits [33]. Our preliminary attempts to apply such models to our song sample have supported Leroi and Swire's [21] claim that musical evolution is much less "tree-like" than genetic evolution, with much independent invention (convergence), and borrowing (horizontal transmission) of individual musical features and even entire songs between populations.

This brings up the more general issue of the dynamics of musical evolutionary change. There are cultural forces that both diversify and homogenize musical repertoires, and some of them are conceptually analogous to forces that influence the dynamics of genetic change [34]. As with genes, cultural forms like songs can undergo random changes over time, a kind of musical "drift" [35]. Small population sizes may enhance the effects of genetic drift, although it is unclear as of yet how population sizes affect musical diversity and change over time. Another major force that can diversify repertoires is admixture through cultural contact, a kind of musical "flow". Recent contact situations, such as between the Paiwan and Rukai of Taiwan in our sample [36], can lead to high levels of acculturation, despite the maintenance of distinct languages. This particular contact situation is well reflected musically, with Paiwan and Rukai producing the only negative pairwise Φ_{ST} value in our analysis. This is unsurprising as music actually provides an excellent model for "hybridization" in the cultural domain, since it is composed of a series of modular components (mainly pitch and rhythm) that can undergo "syncretisms" or blendings of features. A good example of this is found in African-American music, which contains a novel fusion of European tonal features and African rhythmic features [37]. Other cultural forces that can affect the frequency of cultural variants within and between populations include convergence, borrowing, innovation, conformity, extinction, and replacement (through imposition, as in situations of conquest or economic globalization).

One means by which musical repertoires diversify internally is through a fissioning into an increasing numbers of genres or functional song-types, a universal feature of musical repertoires. A classic example of genre-based variation in song structure is found in Arom's work on the music of the Pygmies of the Central African Republic [38], which qualitatively describes systematic differences in the musical features of songs performed in different social contexts, comprising roughly two dozen distinct musical genres (e.g., music for the hunting of elephants, music for the birth of twins). This is likely to be the same case with our Austronesian musical sample. Unfortunately, the limited number of songs in the current study prevented us from doing any sort of meaningful genre-level analysis. It is plausible that some genres of song are less malleable or prone to borrowing, which could affect our results. Given a larger, more comprehensive dataset, the AMOVA approach could be used to explore how variability in genres is structured between and within populations.

Our work on the cultural evolution of music has important limitations, especially as related to our use of archival material. The reliance of our work on archival recordings highlights the difficulty in sampling the musical variation of indigenous populations in the modern world. One concern for the current work is that the kinds of songs represented in the archives that we used did not cover all of the genres of a population's musical repertoire due to ascertainment bias. This, however, does not negate our major finding, as the inclusion of unrecorded music of other genres in our analyses would most likely have increased, not reduced, the internal diversity of the musical repertoires.

Archival recordings are essential in a world where globalization and the associated expansion of Western culture threaten to extinguish much of the rich cultural diversity seen in human populations across the globe [39]. This decline is reflected in the sheer proportion of living languages classified as vulnerable, endangered, or critical, which is at least 27%, according to a conservative recent analysis [40]. The dominant influence of western music has led to non-traditional (western) musical features being incorporated into indigenous musical repertoires through a kind of imposed hybridization. Archival recordings reduce the potential of encountering this form of unwanted admixture but are problematic in other ways.

In addition to the possible sampling bias discussed above, some recordings may be poorly documented, misclassified, non-traditional, or of poor recording quality. We were fortunate enough to work with a very well-documented archive and to have received advice from Ying-Fen Wang, an ethnomusicologist with expertise in the traditional musics of the Taiwan aborigines. This kind of work may be substantially more difficult in regions with less-organized archives and where ethnomusicological expertise on these traditional musics is lacking. Despite the inherent difficulty in doing this kind of work, the task of characterising and comparing worldwide musical diversity, as other scholars have done with languages [4], is an extremely important endeavour, not least considering the current rapid rate of cultural extinction [40].

c) How generalizable are these results to other aspects of culture?

Many useful parallels have been drawn between cultural and biological evolution [41] but the forces shaping cultural diversity can differ markedly from those that drive the structure of genetic diversity [42]. For example, some have argued that cultural variants will necessarily always display less intra-population variation than will genetic variants [11]. Language is one of the best-cited examples of a cultural trait that is mostly variable between speech communities (rather than within), due to strong constraints that ensure that members of a speech community can communicate with one another [10]. The relative strength of processes that reduce internal diversity and those that increase it is likely to differ across cultural domains. It is plausible that music, for example, may be subject to lesser constraints than a system like language, and that innovation in this domain may be more highly valued in some cultures. The current work only covers musical variation in a small number of populations within the same language family. Populations in other regions of the world may have much more homogeneous musical repertoires. However, our results demonstrate that a high degree of internal heterogeneity in a population's musical repertoire is the reality, in at least some cases.

d) Conclusion

While the present-day structure of human genetic diversity has been rigorously quantified, we lack the same kind of quantitative information for most aspects of culture.

The AMOVA framework provides cross-cultural researchers a means of quantifying variability for a number of cultural forms, and of exploring the forces responsible for balancing diversity and conformity. The current work is by no means meant as a comprehensive sampling of worldwide musical variability, and indeed the partitioning of musical variance may differ substantially in other regions of the world. We do, however, present a crucial tool that can be applied to many other aspects of culture, a tool that can be useful for the study of human migrations and associated histories of cultural contact.

ACKNOWLEDGEMENTS

This work was supported by a grant to S.B. from the Social Sciences and Humanities Research Council of Canada and by an Amherst College Roland Wood Fellowship to P.E.S. We would like to thank Tom Currie for helpful comments on an earlier version of this manuscript. We would also like to thank Jean Trejaut, Yingfen Wang and Victor Grauer for advice and support while conducting this research, as well as Michel Belyk for advice on programming.

REFERENCES

1 Lewis, M. (ed) 2009 *Ethnologue: languages of the world*. 16th ed. Dallas, TX: SIL International.

2 Murdock, G. 1967 Ethnographic atlas: A summary. *Ethnology* **6**, 109-236.

3 Haspelmath, M., Dryer, M., Gil, D. & Comrie, B., editors 2005 *The World Atlas of Language Structures*. Oxford: Oxford University Press.

4 Greenhill, S. J., Blust, R. & Gray, R. D. 2008 The Austronesian Basic Vocabulary Database: from bioinformatics to lexomics. *Evol. Bioinform.* **4**, 271-83.

5 Currie, T. E., Greenhill, S. J., Gray, R. D., Hasegawa, T. & Mace, R. 2010 Rise and fall of political complexity in island South-East Asia and the Pacific. *Nature* **467**, 801-804. (doi:10.1038/nature09461)

6 Gray, R. D., Drummond, A. J. & Greenhill, S. J. 2009 Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* **323**, 479-83. (doi:10.1126/science.1166858)

7 Reesink, G., Singer, R. & Dunn, M. 2009 Explaining the linguistic diversity of Sahul using population models. *PLoS. Biol.* **7**, e1000241. (doi:10.1371/journal.pbio.1000241)

8 Lewontin, R. C. 1972 The apportionment of human diversity. In *Evolutionary Biology* (eds. T. Dobzhansky M. K. Hecht & W. Steere), pp. 391-398. New York, N.Y. Appleton-Century-Crofts.

9 Rosenberg, N. a, Pritchard, J. K., Weber, J. L., Cann, H. M., Kidd, K. K., Zhivotovsky, L. a & Feldman, M. W. 2002 Genetic structure of human populations. *Science (New York, N.Y.)* **298**, 2381-5. (doi:10.1126/science.1078311)

10 Newson, L., Richerson, P. J. & Boyd, R. 2007 Cultural Evolution and the Shaping of Cultural Diversity. In *Handbook of Cultural Psychology* (eds. S. Kitayama & D. Cohen), pp. 454-476. New York, N.Y. The Guildford Press.

11 Foley, R. & Mirazón Lahr, M. 2011 The evolution of the diversity of cultures. *Philos. Trans. R. Soc. B* **366**, 1080. (doi:10.1098/rstb.2010.0370)

12 Bell, A. V., Richerson, P. J. & McElreath, R. 2009 Culture rather than genes provides greater scope for the evolution of large-scale human prosociality. *Proc. Natl. Acad. Sci. USA.* **106**, 17671-4. (doi:10.1073/pnas.0903232106)

13 Nettl, B. 2000 An ethnomusicologist contemplates universals in musical sound and musical culture. In *The origins of music* (eds. N. L. Wallin B. Merker & S. Brown), pp. 463–472. Massachusetts: The MIT Press.

14 Lomax, A. 1968 *Folk song style and culture*. New Brunswick, N.J. Transaction Books.

15 Henry, E. 1976 The variety of music in a north Indian village: Reassessing Cantometrics. *Ethnomusicology* **20**, 49-66.

16 Feld, S. 1984 Sound structure as social structure. *Ethnomusicology* **28**, 383–409.

17 Excoffier, L., Smouse, P. & Quattro, J. 1992 Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**, 479.

18 Michalakis, Y. & Excoffier, L. 1996 A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* **142**, 5-8.

19 Bryant, D. & Moulton, V. 2004 Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Mol. Biol. Evol.* **21**, 255-65. (doi:10.1093/molbev/msh018)

20 National Taiwan Preparatory Office of Traditional Arts 2010 Taiwan National Music Archive. http://music.ncfta.gov.tw [in Chinese]

21 Leroi, A. M. & Swire, J. 2006 The recovery of the past. *World of music* **48**, 43-54.

22 Busby, G. 2006 Finding the blues: An investigation into the origins and evolution of African-American music. 1-55.

23 R Development Core Team 2011 R: A Language and Environment for Statistical Computing.

24 Lingoes, J. 1971 Some boundary conditions for a monotone analysis of symmetric matrices. *Psychometrika* **36**, 195-203.

25 Dray, S. & Dufour, A. B. 2007 The ade4 package: implementing the duality diagram for ecologists. *J. Stat. Softw.* **22**, 1–20.

26 Excoffier, L., Laval, G. & Schneider, S. 2005 Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform.* **1**, 47-50.

27 Huson, D. H. & Bryant, D. 2006 Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **23**, 254-67. (doi:10.1093/molbev/msj030)

28 Meirmans, P. G. 2006 Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution* **60**, 2399–2402.

29 Mitton, J. B. 1977 Genetic differentiation of races of man as judged by single-

locus and multilocus Analyses. *Am. Nat.* **111**, 203–212.

30 Mitton, J. B. 1978 Measurement of differentiation : Reply to Lewontin, Powell, and Taylor. *Am. Nat.* **112**, 1142–1144.

31 Edwards, a W. F. 2003 Human genetic diversity: Lewontin's fallacy. *BioEssays* **25**, 798-801. (doi:10.1002/bies.10315)

32 Cavalli-Sforza, L. L., Feldman, M. W., Chen, K. H. & Dornbusch, S. M. 1982 Theory and observation in cultural transmission. *Science* **218**, 19-27.

33 Greenhill, S. J., Currie, T. E. & Gray, R. D. 2009 Does horizontal transmission invalidate cultural phylogenies? *Proc. R. Soc. B.* **276**, 2299-306. (doi:10.1098/rspb.2008.1944)

34 Cavalli-Sforza, L. L. & Feldman, M. W. 1981 *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton, N.J. Princeton University Press.

35 Bentley, R. A., Hahn, M. W. & Shennan, S. J. 2004 Random drift and culture change. *Proc. R. Soc. B.* **271**, 1443-50. (doi:10.1098/rspb.2004.2746)

36 Huteson, G. 2003 Sociolinguistic survey report for the Tona and Maga dialects of the Rukai language. *SIL Electronic Survey Reports* **12**.

37 Kubik, G. 1999 *Africa and the blues*. Jackson, M.S. University Press of Mississippi.

38 Arom, S. 1991 *African Polyphony & Polyrhythm: Musical Structure and Methodology*. Cambridge, U.K. Cambridge University Press.

39 Lomax, A. 1977 Appeal for cultural equity. *J. Commun.* **27**, 125–138.

40 Sutherland, W. J. 2003 Parallel extinction risk and global distribution of languages and species. *Nature* **423**, 276–279. (doi:10.1029/2002GL015811)

41 Henrich, J., Boyd, R. & Richerson, P. J. 2008 Five misunderstandings about cultural evolution. *Hum. Nature* **19**, 119-137. (doi:10.1007/s12110-008-9037-1)

42 Mesoudi, A., Whiten, A. & Laland, K. N. 2006 Towards a unified science of cultural evolution. *Behav. Brain. Sci.* **29**, 329-47; discussion 347-83. (doi:10.1017/S0140525X06009083)

3.6 Supplementary Material

S1 – "CantoCore" Musical Classification Scheme

I) "MELODY" (between-note)

A) Rhythm

1) METER

Cyclic groupings of strong and weak beats into bars

(a) *A-metric:* No consistent beat
(b) *Hetero-metric:* There is a consistent beat, but strong and weak beats occur without a consistent pattern
(c) *Poly-metric*:* Multiple cyclic patterns of strong and weak beats coexist simultaneously (e.g., 6/8 against 3/4)
(d) *Iso-metric:* There is a single, consistent pattern of strong and weak beats (e.g., 3/4, 6/8, 5/4, 2+2+3/8)

N.B. Songs not classified as "*iso-metric*" must be coded "?" for parameters **(2-5)**.

2) NUMBER OF BEATS

The number of beats in a bar

(i) *Duple:* Multiples of 2 (e.g., 4/4, 6/8, 2+3/8)
(ii) *Triple:* Multiples of 3 (e.g., 3/4, 9/8, 2+2+3/8)
(iii) *Complex:* Multiples of prime numbers greater than 3 (e.g., 5/4, 5/8, 2+2+3+2+3/8)
(?) *A-/hetero-/poly-metric:* See **(1)**

3) BEAT SUB-DIVISION

Division of beats into sub-beat-level metric groupings

(a) *A-divisive:* Beats are not sub-divided (e.g., a 4/4 piece containing only ♩ and 𝅗𝅥 notes)
(b) *Hetero-divisive:* Beats are sub-divided, but the number of sub-beats per beat changes (e.g., 2+2+3/8)
(c) *Iso-divisive:* Beats sub-divided into a consistent number of sub-beats (e.g., 6/8, a 4/4 piece containing ♪ notes)
(?) *A-/hetero-/poly-metric:* See **(1)**

N.B. Songs not classified as "*iso-divisive*" must be coded "?" for parameter **(4)**.

4) NUMBER OF SUB-BEATS

The number of sub-beats in a beat

(i) *Simple:* Multiples of 2 (e.g., ♩ beat divided into ♪ note sub-beats; includes 3/4, 4/4, etc.)

(ii) *Compound:* Multiples of 3 (e.g., ♩. beat divided into ♪ note sub-beats; includes 6/8, 9/8, "swing", etc.)

(iii) *Complex:* Multiples of prime numbers greater than 3 (e.g., ♩ beat divided into 5 sub-beats)

(?) *A-/hetero-/poly-metric or a-/hetero-/poly-divisive:* See **(1/3)**

N.B. Songs not classified as "*iso-divisive*" must be coded "?" for parameter **(4)**.

5) SYNCOPATION

The percentage of notes that are accented but in a metrically weak position

(i) *Un-syncopated:* <5%
(ii) *Moderately syncopated: 5-20%*
(iii) *Highly syncopated:* >20%
(?) *A-/hetero-/poly-metric:* See **(1)**

6) MOTIVIC REDUNDANCY

The percentage of all notes that are constructed from a single recurring rhythmic pattern

(i) *Non-motivic:* <20%
(ii) *Moderately motivic: 20-50%*
(iii) *Highly motivic*: >50%

7) DURATIONAL VARIABILITY

Maximum number of different types of duration values in a song

(i) *Low durational variability:* <3 duration values (e.g., only ♪ and ♩)

(ii) *Moderate durational variability:* 3-4 duration values (e.g., ♪, ♩ and 𝅗𝅥)

(iii) *High durational variability:* >4 duration values (e.g., 𝅘𝅥𝅯, ♪, ♩., 𝅗𝅥 and 𝅝)

B) Pitch

8) TONALITY

Organization of discrete pitches around one or more tonic notes

(a) *Indeterminate a-tonal:* No discrete pitches (e.g., exclamations, heightened speech)
(b) *Discrete a-tonal:* Discrete pitches, but no tonic
(c) *Hetero-tonal:* Tonic modulates/shifts between phrases
(d) *Poly-tonal:* Multiple, simultaneous tonics in different vocal parts
(e) *Iso-tonal:* Single tonic throughout

N.B. Songs not classified as *"iso-tonal"* must be coded "?" for parameters **(9-11)**.

9) MODE

Presence of scale degrees at a minor 3rd (250-350 cents) or major 3rd (350-450 cents) above the tonic

(a) *A-modal:* No 3rd present
(b) *Hetero-modal:* Both major and minor 3rd appear in separate phrases
(c) *Poly-modal:* Both major and minor 3rd appear in the same phrase
(d) *Minor iso-modal:* Minor 3rd only
(e) *Major iso-modal:* Major 3rd only
(?) *A-/hetero-/poly-tonal:* See **(8)**

10) NUMBER OF SCALE DEGREES

Number of scale degrees found in the scale

(i) *Sparse scale:* <4 scale degrees
(ii) *Moderately dense scale:* 4-5 scale degrees
(iii) *Dense scale:* >5 scale degrees
(?) *A-/hetero-/poly-tonal:* See **(8)**

11) HEMITONICITY

Frequency of melodic intervals that are semitones (50-150 cent intervals)

(i) *Anhemitonic:* <5%
(ii) *Moderately hemitonic:* 5-20%
(iii) *Highly hemitonic:* >20%

12) MELODIC INTERVAL SIZE

Maximum pitch distance between successive notes within any vocal part

(i) *Small intervals:* <350 cents (i.e., minor 3rd or less; formerly divided into "*monotone*", *"narrow"*, and "*diatonic"* intervals)
(ii) *Medium intervals:* 350-750 cents (i.e., major 3rd - perfect 5th; formerly divided into "*wide*" and "*very wide*" intervals)
(iii) *Large intervals*:* >750 cents (i.e., minor 6th or greater)

13) MELODIC RANGE

Maximum pitch distance between the highest and lowest notes within any vocal part

(i) *Small range:* <750 cents (i.e., perfect 5th or less)
(ii) *Medium range:* 750-1250 cents (i.e., perfect 5th - octave)
(iii) *Large range:* >1250 cents (i.e., more than an octave)

14) MELODIC CONTOUR

Shape resulting from all changes in interval direction within a vocal part

(a) *Horizontal*:* No ascending or descending intervals
(b) *Ascending*:* Ascending intervals only
(c) *Descending:* Descending intervals only

(d) *U-shaped*:* First descending, then ascending intervals
(e) *Arched:* First ascending, then descending intervals
(f) *Undulating:* Multiple changes of interval direction

C) Syllable

15) MELISMA

Maximum number of consecutive notes without articulating a new syllable

(i) *Syllabic:* 1-2 notes
(ii) *Mildly melismatic:* 3-5 notes
(iii) *Strongly melismatic:* >5 notes

16) VOCABLES

The percentage of syllables containing only vowels or semi-vowels (e.g., "y", "h", "w")

(i) *Few vocables:* <20%
(ii) *Some vocables:* 20-50%
(iii) *Many vocables:* >50%

II) "TEXTURE" (between-part)

17) NUMBER OF VOCAL PARTS

Maximum number of simultaneous vocal parts

(i) *One-part:* 1 (formerly divided into "*solo*" and "*unison*")
(ii) *Two-part*:* 2
(iii) *Many-part*:* >2

N.B. Songs classified as "*one-part*" (including both unison and solo songs) must be coded "?" for parameters **(18-20)**.

18) RHYTHMIC TEXTURE (Cantometrics Line 12)

Temporal asynchrony in the relative onsets of different vocal parts (in seconds)

(a) *Hetero-rhythmic (heterophonic):* 0.1–1s
(b) *Poly-rhythmic (polyphonic):* >1s
(c) *Iso-rhythmic (homophonic):* <0.1s
(?) *One-part (monophonic):* See **(17)**

N.B. Songs not classified as "*iso-rhythmic*" must be coded "?" for parameter **(19)**.

19) HARMONIC TEXTURE

Minimum harmonic interval (octave-equalized) between simultaneous vocal parts that is sustained for at least 1 second

(i) *Dissonant:* 50-250 cents (e.g., $2^{nds}/7^{ths}$)
(ii) *Consonant:*>250 cents (e.g., 3^{rds}-6^{ths})
(?) *One-part, or poly-/hetero-rhythmic:* See **(17/18)**

20) RELATIVE MOTION

Relationship of the melodic contours (see **13**) of two simultaneous parts

(d) *Hetero-contour (drone):* One part is horizontal, the other changes direction
(e) *Poly-contour (independent motion):* Both parts have different, non-horizontal contours
(f) *Iso-contour (parallel motion):* Both parts have the same contour
(?) *One-part: S*ee **(17)**

III) "FORM" (between-phrase)

21) PHRASE REPETITION

Maximum number of successive phrases before a phrase is repeated

(i) *Non-repetitive*: >8 phrases, or no repeat at all
(ii) *Moderately repetitive:* 3-8 phrases
(iii) *Repetitive:* 1-2 phrases

22) PHRASE LENGTH

Maximum phrase length, in seconds

(i) *Short phrases:* <5 s
(ii) *Medium-length phrases:* 5-9 s
(iii) *Long phrases*: >9 s

23) PHRASE SYMMETRY

Ratio of the length of the longest phrase in a song relative to the shortest phrase

(i) *Symmetric:* <1.5 times the length of the shortest phrase
(ii) *Mildly asymmetric*:* 1.5-2.5 times the length of the shortest phrase
(iii) *Very asymmetric*: >2.5 times the length of the shortest phrase

24) SOLO/GROUP ARRANGEMENT

Number of singers in each phrase

(e) *Solo:* Only solo phrases throughout
(f) *Mixed:* Individual phrases contain both group and solo sub-sections
(g) *Alternating:* Alternation between distinct solo and group phrases
(h) *Group:* Only group phrases throughout

25) RESPONSORIAL ARRANGEMENT

Alternation of phrases between different vocal parts

(d) *A-responsorial:* No alternation between parts
(e) *Hetero-responsorial*:* Irregular alternation between parts
(f) *Iso-responsorial:* Consistent alternation between parts

N.B. Songs classified as "*a-responsorial*" must be coded "?" for parameter **(26)**.
Comments: See comments in **(24)**.

26) PHRASE OVERLAP

Maximum overlap between a "call" phrase and the "response" phrase that alternates with it (as the percentage of time in which the latter phrase overlaps with the former)

(i) *Non-overlapping:* 0%
(ii) *Mildly overlapping:* 1–25%
(iii) *Highly overlapping:* >25%
(?) *A-responsorial:* See **(25)**

S2 – Song-to-song distance matrix algorithm description

a) Preparation of data

The non-independent nature of some of the characters in the CantoCore scheme results in the potential for redundancy to be introduced into the codings. For example, if the character "meter" is coded as "a-metric" (i.e., no recognisable beat), then the codings for the characters "number of beats", "sub-beat division", and "number of sub-beats" carry the same information, and this can overemphasise the importance of the absence of a beat. In order to reduce this potential redundancy, a "?" is put in place of redundant codings and is treated as missing data. Uncoded characters are also denoted with a "?".

b) Ordinal and nominal variables

CantoCore contains both ordinal and nominal characters. The first part of the distance-matrix algorithm separates ordinal and nominal variables, which are defined a priori by the user. Ordinal characters are coded using lower-case Roman numerals (i-iii in the case of CantoCore) and are never multi-coded. Nominal characters are coded using lower-case letters (a-f in the case of CantoCore), and can be multi-coded where appropriate.

c) Ordinal characters

Ordinal characters can have a number of possible character states, and this can vary across characters. In order to keep this consistent across ordinal characters, as well as for ordinal comparisons to be equivalent to nominal comparisons, the raw codings of the ordinal character states are converted to *scaled* values from 0 to 1 such that the minimum ordinal value is coded as 0 and the maximum ordinal value is coded as 1, with intermediate states taking intermediate values (see figure S1 a-b). In CantoCore, all ordinal characters have three possible states (i, ii, or iii). As a result, i becomes 0, ii becomes 0.5, and iii becomes 1. For a character with five states (i, ii, iii, iv, and v), the converted values would become 0, 0.25, 0.50, 0.75, and 1, respectively.

The algorithm then creates a separate pairwise distance matrix for each ordinal character by taking the absolute difference between the scaled codings for each pair of songs (Figure S1 b-c):

$$| \text{Song X} - \text{SongY} |$$

The maximum possible difference is 1 (if the codings are maximally different) and the minimum value is 0 (if the codings are identical). For pairs of songs where one or both songs lack codings for that character ("?"), the distance is listed as NA (See Figure S1 c). The end result is a separate pairwise distance matrix for each ordinal character, from 1-j.

(a)

Song	Ordinal variable j
Song A	iii
Song B	i
Song C	?
Song D	ii

i = 0
ii = 0.5
iii = 1
? = ?

Codings scaled from 0-1

(b)

Song	Ordinal variable j
Song A	1
Song B	0
Song C	?
Song D	0.5

Distance calculated as
| Song X – Song Y |

(c)

j	Song A	Song B	Song C	Song D
Song A	0	1	NA	0.5
Song B	1	0	NA	0.5
Song C	NA	NA	NA	NA
Song D	0.5	0.5	NA	0

Figure S1. A sample calculation of pairwise distance across 4 songs (A-D) for a single ordinal variable (j). (a) The raw CantoCore codings. (b) The same codings after they have been converted into a scale from 0 to 1. (c) A sample distance matrix based on the absolute difference of the scaled codings for each pair of songs. Any pairwise distance involving a redundant or uncoded character (?) is denoted as NA. This overall process is repeated for each ordinal character from 1-j, resulting in j distance matrices for the song set.

d) Nominal characters

In order to accommodate the potential for multi-coded characters, the matrix of raw codings (comprised of lower-case letters) is converted into a "presence-absence" matrix, as in Busby (2006), where Y denotes the presence of a character state, and N denotes its absence (see Figure S2 a-b). Character states coded as "?" (representing redundant or missing codings) are denoted as NA when they are converted into the presence-absence matrix and are thus treated as missing data.

The conversion of scores in the presence-absence matrix into distances is based on pairwise matching of songs across all possible character states. For example, if one song contains an “a” coding (Y for character state a) and another song lacks it (N for character state a), then the program scores this as a 1, implying maximum distance between the two songs. If both songs contain a “b” coding (i.e., both are scored as Y for character state b), then the program scores this as a 0, implying minimum distance between them.

(a)

Song	Nominal variable k
Song A	ad
Song B	c
Song C	?
Song D	d

Conversion of raw codings into a presence-absence matrix

Y = Present
N = Absent
NA = ?

(b)

	Character states			
Var. k	a	b	c	d
Song A	Y	N	N	Y
Song B	N	N	Y	N
Song C	NA	NA	NA	NA
Song D	N	N	N	Y

Calculation of mean pairwise distance across all character states

If both Y, then = 0
If one Y and other N = 1
If both N, then don’t include in mean
If one or both NA, then denote NA

(c)

k	Song A	Song B	Song C	Song D
Song A	0	(1+1+1)/3=1	NA	(1+0)/2=0.5
Song B	(1+1+1)/3=1	0	NA	(1+1)/2=1
Song C	NA	NA	NA	NA
Song D	(1+0)/2=0.5	(1+1)/2=1	NA	0

Figure S2. A sample calculation of mean distance for a nominal variable. Raw CantoCore codings (a) are first converted into a “presence-absence” matrix (b), where Y denotes the presence of a character state, N denotes its absence, and NA represents a redundant coding for each song. The mean pairwise distance is calculated by taking the pairwise distance between songs across all character states, except those involving mutual absence, which are ignored. This process is repeated for each nominal variable from 1-k.

If both songs lack a particular character state, for example c for Songs A and D (where both are scored as N), then this mutual absence is ignored and is not incorporated into the mean distance calculation. This is done because the mutual absence of a character state is uninformative.

Finally, to calculate the mean pairwise distance for a particular character, we take the pairwise distances between songs across all character states for that character, except those involving mutual absence, which are ignored. The occurrence of mutual absence results in some distances (means) containing fewer comparisons than others. As a result, the denominator in the mean calculation is variable. For example, the AB distance contains three comparisons while the AD distance contains only two, since the latter pair has two mutual absences compared to only one for the former pair. As for the ordinal characters, 1 is the maximum possible mean distance, and 0 is the minimum. If one or both of the songs of a pair contain NA's anywhere in their fields (because the raw coding was a "?", as with Song C), the pairwise distance is denoted as NA. A separate distance matrix is created this way for each nominal variable from 1-k.

e) Combining ordinal and nominal characters into a final distance measure

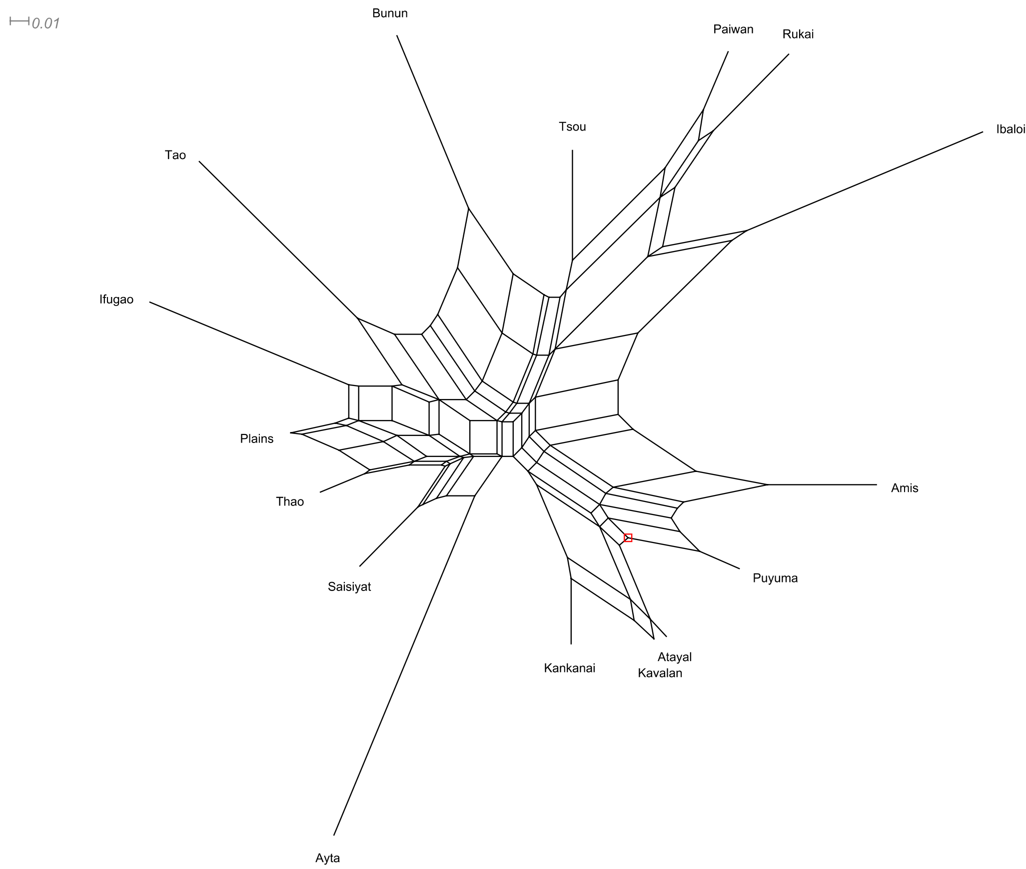
The final step of the algorithm combines information from the j ordinal variables with the k nominal variables to obtain an overall measure of distance between songs. For each pair of songs, the mean distance across all characters is taken, ignoring any distances denoted as NA. As a result, the final measure of distance incorporates information from each character equally, ignoring only redundant codings or uncoded characters.

S3 – Modal Profiles

Population	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
Amis	d	a	c	a	0.5	1	0	e	d	0.5	0	0.5	1	f	0	1	0	b	?	b	1	1	0.5	c	c	0.5
Atayal	d	a	c	a	0	1	0	e	d	0.5	0	0.5	0.5	f	0	0	0	?	?	?	0.5	0.5	0	d	a	?
Ayta	a	?	?	?	?	0.5	0	e	a	0.5	0	0.5	0.5	c	0	0.5	0	?	?	?	1	0.5	0	c	c	0.5
Bunun	a	a	?	?	0	1	0.5	e	e	0.5	0	0.5	1	a	0.5	0	1	c	0.5	ca	1	1	0	b	a	?
Ibaloi	a	?	?	?	?	0.5	1	e	e	0.5	1	0.5	0.5	fe	0.5	0.5	0	?	?	?	1	1	1	c	c	0
Ifugao	a	?	?	?	?	0.5	0.5	e	?	?	0	0.5	0	f	1	0.5	1	ac	0	cb	1	0.5	0.5	d	a	?
Kankanai	d	?	?	?	?	1	0	e	e	0.5	0	0.5	0.5	f	0	0	0	?	?	?	1	0.5	0.5	d	a	?
Kavalan	d	a	c	a	0	1	0	e	d	0.5	0	0.5	0.5	f	0	0	0	?	?	?	0.5	0.5	0.5	d	a	?
Paiwan	b	?	?	?	?	1	0	e	e	0.5	0	0.5	0.5	f	0.5	0	0.5	c	0.5	a	1	1	0.5	c	a	?
Plains	a	?	?	?	?	0.5	0.5	e	d	0.5	0	0.5	1	f	0.5	0.5	0	?	?	?	1	0.5	0	d	a	?
Puyuma	d	a	c	a	0	1	0	e	d	0.5	0	0.5	1	f	0.5	1	0	?	?	?	0.5	1	0.5	d	a	?
Rukai	b	?	?	?	?	1	0.5	e	e	0.5	1	0.5	0.5	fa	0.5	0	0	?	?	?	1	1	0.5	cb	a	?
Saisiyat	a	?	?	?	?	1	0	e	d	0	0	0.5	0.5	f	0.5	0.5	0	?	?	?	1	0.5	0	c	a	?
Tao	a	?	?	?	?	0.5	1	e	d	0	0	0.5	0	a	0	0	0	?	?	?	1	1	0	d	a	?
Thao	a	?	?	?	?	1	0	e	d	0.5	0	0.5	1	f	0.5	0.5	0	?	?	?	0.5	0.5	0	d	a	?
Tsou	d	?	?	?	?	1	0	e	e	0.5	0	0.5	1	f	0.5	0.5	0.5	c	0.5	c	1	1	0.5	c	a	?

We created modal profiles by taking the most commonly coded character state for every CantoCore character in each population's musical repertoire. Numbers across the top correspond to the 26 CantoCore categories found in S1. Nominal character states are coded as letters and multiple states are permitted. Ordinal characters are coded as numbers. "?" codings are treated as missing data, because their inclusion would carry information redundant with the coding of another character.

S4 – Neighbor-Net of modal profile distances



Chapter 4

Coevolution of music and genes in aboriginal Taiwan

Patrick E. Savage[1], Tom Rzeszutek[1], Victor Grauer[2], Ying-fen Wang[3], Jean Trejaut[4], Marie Lin[4], and Steven Brown[1]

[1]Department of Psychology, Neuroscience & Behaviour, McMaster University, Canada
[2] Independent scholar, Pittsburgh, USA
[3]Graduate Institute of Musicology, National Taiwan University, Taiwan
[4]Transfusion Medicine Laboratory, Mackay Memorial Hospital, Taiwan

Keywords: music; genetics; evolution; migration; Cantometrics

4.1 Abstract

A recent claim that music can act as a marker of ancient population migrations has been criticized on the grounds that music's time-depth seems too shallow (i.e., music changes too rapidly to preserve deep relationships). We predicted that, if any musical features were to have the necessary time-depth, they would be the structural features – rather than performance features – of group songs. To test this prediction, we used Cantometric classifications of 234 traditional group songs from 9 Taiwanese aboriginal tribes to create separate distance matrices for music based on either structural or performance features. Both distance matrices demonstrated a positive correlation with distances based on mitochondrial DNA, a migration marker with well-established time-depth. However, this correlation was only statistically significant for the distance matrix based on song structure, which accounted for more than twice as much variance in the genetic data as did performance style. These results, although preliminary, provide the first quantitative evidence suggesting that songs and genes may have followed parallel evolutionary trajectories on the order of several thousand years. Follow-up analyses using different coders, different classification schemes, and different genetic data-sets, while partially confirming these patterns, highlight substantial challenges in obtaining and interpreting reliable samples for both music and genes.

4.2 Introduction

The discovery that our genes can be used to trace the migration of all anatomically modern humans back to a single African "mitochondrial Eve" has had an enormous impact on our understanding of human pre-history (Cann, Stoneking, & Wilson, 1987). Recently, Grauer (2006) proposed that music, too, traces these same migrations back to Africa, but critics argued that music changes too rapidly to preserve ancient relationships (e.g., Stock, 2006). These opposing claims present an excellent opportunity to bring new quantitative techniques and data from genetic anthropology to old questions of how and why music varies cross-culturally.

Grauer's claim drew on data from the landmark Cantometrics Project (Lomax, 1968), which remains the only global scientific study of human song. At the time, Lomax's causal interpretation of the correlation between culture and music – for example, male dominance causing nasal singing – was highly criticized even by other members of the Cantometrics Project (e.g., Erickson, 1976). While Grauer's recent migratory interpretation avoids Lomax's pitfall, Stock's rebuttal still echoed many of the original criticisms of the Cantometrics Project (e.g., Blacking, 1977; Downey, 1970). Could the acoustic surface of music really reflect ancient connections between cultures? If so, are these reflected in performance features, or in the structural features traditionally emphasized in Western musicology?

Lomax himself was highly critical of the use of Western musical notation in ethnomusicology, which he saw as emphasizing surface structural features at the expense of deeper performance features. He spent his life developing a performance-oriented approach that was concerned "not with songs abstracted from the stream of vocalizing we encountered on the tapes, but with the stream itself, with 'singing'" (Lomax, 1980). Nevertheless, the Cantometric classification scheme that Lomax and Grauer (1968) developed contained roughly equal numbers of features devoted to song structure and singing style (see Fig. 4.1).

Our own view differs from both Lomax's and his critics' in that we propose that the structural features of song should have the greatest time-depth to track migrations, especially when applied to group songs. Our reasoning is that structural features such as melody, texture and form require greater consensus among singers than the more idiosyncratic variation that goes into performance, such as timbre or ornamentation. Hence, features like scales and rhythms should be more stable over time than features like nasality or rubato.

These claims are testable. As a case-study to examine music's time-depth in the context of human migrations, we have examined the traditional group songs of the aboriginal tribes of Taiwan, who have been well-studied in terms of music, genetics, linguistics, and migration (Diamond, 2000; Kurosawa, 1973; Trejaut et al., 2005). Our primary aim, therefore, was to use existing information about the relative patterns of genetic and musical similarity among the Taiwanese aboriginal tribes to empirically test for the first time whether song structure or singing style has the time-depth required for studying human migrations. Our basic method was to compare music – a marker of unknown time-depth – against the best available marker with a well-established time-depth, namely mitochondrial DNA (mtDNA) (Oppenheimer, 2004).

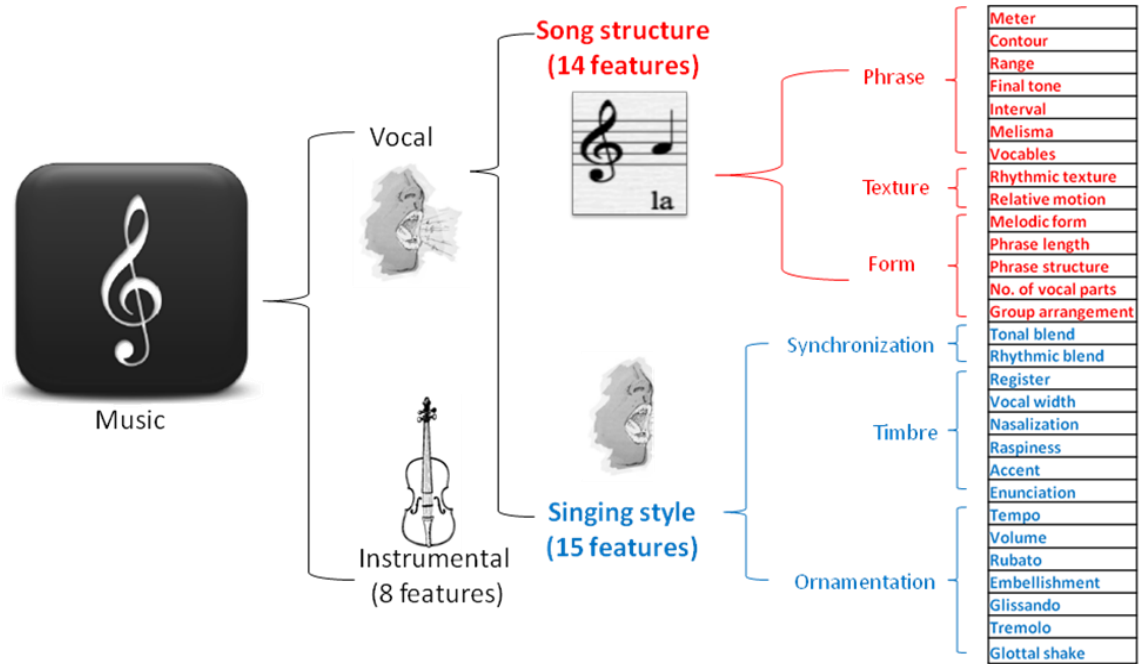


Fig. 4.1: **Organization of the 14 song-structure (red) and 15 singing-style (blue) Cantometric classification features used in this analysis. Note that our method focuses on the vocal component of the music and therefore ignores 8 classification features related to instruments.**

4.3 Methods

Genetic sample: Genetic samples were obtained from JT and ML's previously published data (Trejaut et al., 2005). These samples consist of hypervariable segments 1 and 2 in the control region of the maternally inherited, haploid mitochondrial genome from 640 individuals. After editing, a 717-base-pair (bp) nucleotide string was used in the current analysis, with the same sample sizes previously reported.

Musical sample: Using the Taiwan National Music Archive[1], we obtained 234 traditional group songs from all 9 of the tribes whose mtDNA data was published in Trejaut et al. (2005). For populations with 30 or fewer eligible songs, all eligible songs in the archive were used. For populations with over 30 eligible songs, we randomly sampled 30 songs per population. Sample sizes were: Amis=30, Atayal=8, Bunun=30, Paiwan=30, Puyuma=30, Rukai=30, Saisiyat=26, Tao (Yami)=29, Tsou=21.

Distances between samples: Pairwise distances between individual genetic samples were calculated based on the number of pairwise differences between nucleotide sequences. This is the simplest possible measure of genetic distance, as it is based on phenetic (similarity) relationships rather than phylogenetic (evolutionary) relationships that require additional assumptions about how differences arose. Building on the work of Leroi and

[1] http://music.ncfta.gov.tw (in Chinese)

Swire (2006) and Busby (2006), we developed a comparable phenetic distance measure for Cantometrics that accounts for various statistical complications inherent in converting musical classification schemes into quantitative distances (for details, see Rzeszutek, Savage and Brown, submitted; see Chapter 3 of this thesis). Cantometric classifications of the songs were coded by VG. Two separate musical distance-matrices were calculated: one using the 14 song-structure characters from Cantometrics, the other using the 15 singing-style characters (see Fig. 4.1 for details about these features). Eight Cantometric characters related to instruments alone were excluded from this analysis.

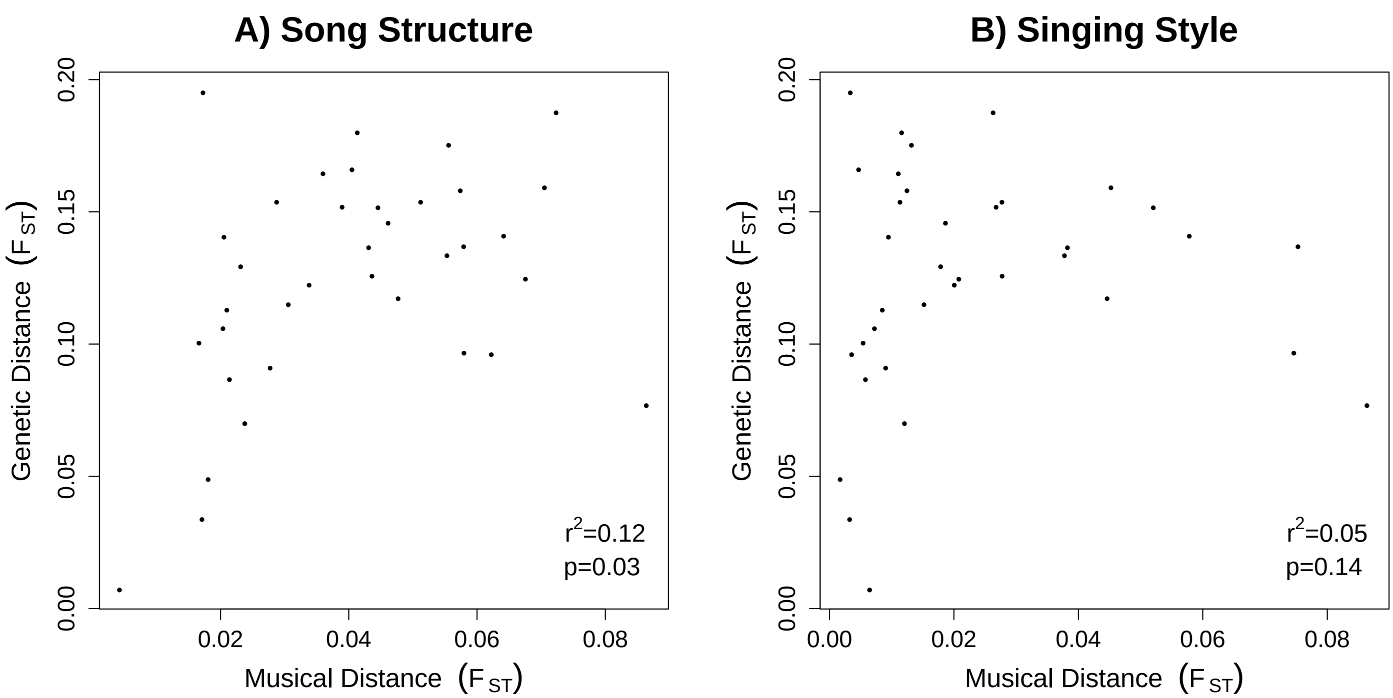
Distances between populations: For both the genetic and musical data, the 36 possible pairwise distances among the 9 tribes were calculated using the Analysis of Molecular Variance (AMOVA) framework (Excoffier, Smouse, & Quattro, 1992), as in Rzeszutek et al. (submitted; see Chapter 3 of this thesis). These distances were measured using a statistic called ϕ_{ST}, which represents the proportion of variability among individual samples that is due to between-population differences. Thus, it explicitly incorporates within-population heterogeneity, avoiding the assumptions of within-population homogeneity that plagued Lomax's original statistical methodology (e.g., Henry, 1976; Leroi & Swire, 2006).

Correlations: The statistical significance of the correlations between musical and genetic distances was tested using the permutation-based Mantel test using 20,000 permutations and Spearman's rho, with the threshold for significance set at $p < 0.05$ (one-tailed). This test controls for the fact that the 36 pairwise distances among the nine tribes are not independent of one another. We used a rank-order Spearman correlation due to the distance nature of our musical and genetic datasets.

Follow-up analyses: In order to explore the effects of using different samples, different classification schemes, and different coders, we also applied the above methodology to the same set of songs coded by PES using an alternative classification scheme, CantoCore (Savage, Merritt, Rzeszutek, and Brown, submitted; see Chapter 2 of this thesis), which is based on Cantometrics, but only contains features related to song structure. Second, we compared both VG's and PES's codings with those of a third coder, Emily Merritt, who was blind to the tribal affiliations of the songs. Inter-rater reliability values were calculated using κ as in Savage et al. (submitted) for a pseudo-randomly selected subset of 45 songs (5 songs randomly selected per tribe). Finally, we obtained an unpublished high-resolution dataset from Albert Ko, Fred Delfin, and Mark Stoneking, containing whole-genome mtDNA samples from 407 individuals from the same nine Taiwanese populations and from 90 individuals from three aboriginal Philippines tribes (Ibaloi, Ifugao, and Kankanai) whose songs are analyzed in Rzeszutek et al. (submitted; see Chapter 3 of this thesis). We performed the same analyses as above with these new data, once using the same 717-bp sub-section of the mtDNA genome as we used for Trejaut et al.'s data, and again using the whole 16,569-bp mtDNA genome.

4.4 Results

The correlation between population-level musical distances based on VG's Cantometric codings and those based on Trejaut et al.'s published genetic data was just beyond the threshold for significance (r^2=0.09, p=0.07). In accordance with our predictions, this trend appeared to be driven by the songs' structural features, as the correlation based on separate distance matrices for structural vs. performance features became significant, but only for features of song structure, which accounted for more than twice as much variance in genetic distance as did features of singing style (structure: r^2=0.12, p=0.03; performance: r^2=0.05, p=0.14) (Figure 4.2). Further analysis using the CantoCore coding scheme, based exclusively on features of song structure that were coded by a different rater, was also significantly correlated with Trejaut's data (r^2=0.10, p=0.045), replicating the above result with the structural features of Cantometrics. Both classification schemes demonstrated similar reliability (Cantometrics: mean κ = 0.28; CantoCore: mean κ = 0.30) that was significantly greater than chance ($p<1x10^{-5}$) but not significantly different from one another (p=0.68). Cantometrics' reliability was similar to



the previously reported value of κ = 0.26, (Savage et al., submitted), while CantoCore was slightly lower than the previously reported value of κ = 0.47. This is probably an effect of the order of training, as EM first classified all Taiwanese songs using CantoCore, then classified them all again using Cantometrics, and then finally classified all of the songs reported in Savage et al. (submitted) using both classification systems at the same time.

Fig. 4.2: **Scatterplots of the 36 pairwise musical and genetic distances among 9 Taiwanese aboriginal tribes. Genetic distances (y-axis) are based on an Analysis of Molecular Variance (AMOVA) of 640 mitochondrial DNA haplotypes from Trejaut et al. (2005). Analogous musical distances (x-axis) were calculated from 234 traditional choral songs using Cantometric characters of either A) song structure or B) singing style (i.e., performance).**

Statistical significance of distance-matrix correlations is based on a Mantel test and Spearman's rho.

Follow-up analyses using the new Taiwanese genetic data-sets were less conclusive. The comparison between VG's Cantometric codings and Ko et al.'s alternative genetic sample gave almost identical r^2 values, although slight changes in the p-values pushed the correlation with structural features just barely beyond the threshold for significance (structure: $r^2=0.12$, $p=0.051$; performance: $r^2=0.03$, $p=0.22$). Surprisingly, however, this replication only held for Ko et al.'s data when we restricted the analysis to the same 717-bp sub-section of the mtDNA genome as was published by Trejaut et al. (2005). When we reran the analysis of Ko et al.'s data using the entire mtDNA genome, the correlations between musical and genetic similarities dropped dramatically (structure: $r^2=0.04$, $p=0.21$; performance: $r^2=0.0001$, $p=0.53$). Although the same pattern of structural features being more strongly correlated with performance features was maintained, both correlations were now nowhere near statistical significance. Also surprising, the correlation between PES's CantoCore codings and the new genetic data-set were substantially reduced when using the new partial-genome data as well as the whole-genome data (partial: $r^2=0.05$, $p=0.16$; full: $r^2=0.03$, $p=0.24$).

Finally, introducing the new data from the three Philippines groups seemed to completely remove any trace of a significant correlation. When these groups were added, all possible permutations of musical features, classification scheme, genetic sample and mtDNA section gave similar, non-significant results ($r^2<0.03$, $p>0.18$).

4.5 Discussion

Our main finding was that musical similarities based on structural features among the 9 Taiwanese aboriginal tribes were significantly correlated with genetic similarities. This provides the first empirical support for Grauer's (2006) claim that music may have the time-depth required for use as a marker in studying prehistoric human migrations. Consistent with our predictions, performance features did not reflect the genetic relationships between populations as strongly as structural features did. The simplest interpretation is that the way a song is performed may be a more rapidly changing or malleable that the underlying structure of the song, which may be relatively more constrained. Similar dissociations in evolutionary rates among different features is also found in linguistics and genetics, where some words are more resistant to borrowing than others (McMahon & McMahon, 2005) and mutation rates vary across the human genome (Nachman & Cromwell 2000).

This does not mean that performance style is not important. The strongest correlations that Lomax originally found were between performance features and social structure (such as vocal tension correlating with sexual restrictiveness), leading him to conclude that structural features such as melody, meter and harmony "…are not connected to the prime characterizers of social and cultural structures" (Lomax, 1980:52). Lomax was fully aware of the complicated interrelationship between music, social structure, and migration, but was never able to resolve debates about the causality of these relationships. Perhaps the differential transmission of structural and performance features is such that song structure better reflects ancient population migrations, while singing style is a better marker of social structure and/or more recent patterns of movement and cultural contact that occurred in the absence of substantial gene flow. Furthermore, performance features may act as a stronger marker of migration in other populations outside of Taiwan.

Our follow-up analyses highlight many methodological challenges in quantifying and interpreting relationships between music and culture. Concerns about reliability, sampling, and the coarse nature of the classification scheme long dogged Cantometrics (Dowling & Harwood, 1986; Maranda, 1970; Nettl, 1970). Although our partial replications using a different classifying scheme, different coder and different genetic sample suggest that some of these effects on analyses are minor, the surprising dissociation between the correlations with the whole mtDNA genome and the partial mtDNA genome suggest that even in population genetics, a coarse analysis can substantially affect the data. Unfortunately, to our knowledge, there is currently no higher-resolution alternative to Cantometrics or CantoCore that is cross-culturally appropriate. This may change soon, however, if the new field of "computational ethnomusicology" (Tzanetakis, Kapur, Schloss & Wright, 2007) begins using these systems as a theoretical framework to build automated algorithms for cross-cultural classification and analysis.

The complete disappearance of correlations when we included the Philippines samples is difficult to interpret. On the one hand, it may imply that music does not function as a marker of prehistoric population relationships in the Philippines, perhaps due to different colonial and political histories. However, it may also reflect differences in the musical samples themselves. We were fortunate to be able to collaborate with a Taiwanese ethnomusicologist (YW) with expertise in traditional aboriginal music using a well-documented Taiwanese archive containing extensive liner notes, many of which were also published commercially. On the other hand, most of the recordings in the Philippines archive are unpublished field recordings containing long, continuous recordings of performances interspersed with talking, laughter, eating, etc. We had no ethnographic notes or ethnomusicological expertise to rely on regarding which songs were representative, which songs were repeat performances of the same song, etc., and thus had to simply use whatever recordings we could find. Our inability to interpret our Philippines data thus highlights the importance of musical sampling and the need for collaboration with expert ethnomusicologists.

Further work is needed to determine the extent to which the genetic and musical connections between the tribes are due to recent admixture through "isolation by distance" (Wright, 1943) as opposed to parallel coevolutionary isolation and drift through "branching" (Cavalli-Sforza, 1997) after diverging from one or more founding populations. For example, while the Paiwan and Rukai languages are mutually unintelligible due to differences that have evolved in the ~6,000 years since their ancestral language, proto-Austronesian, first arrived on Taiwan (Gray, Drummond, & Greenhill, 2009), our analyses found that they are not significantly different from each other either musically or genetically. It appears that musical and genetic similarities between the two tribes are a recent development due to extensive intermarriage and musical exchange, with linguistic differences maintained through bilingualism (Huteson, 2003). For example, the type of drone polyphony that is now so distinctive of both tribes was performed only by the Rukai before World War II (Kurosawa, 1973).

Regardless of how and when the similarities arose, however, our findings in Taiwan lend provisional support for the ability of musical structure to track population movements in the same manner as do genes. Whether the coevolution and co-migration of music and genes extends as far back as Grauer's Out-of-Africa claim, however, remains an open empirical question.

ACKNOWLEDGMENTS

We would like to thank Albert Ko, Fred Delfin, and Mark Stoneking for sharing their unpublished genetic data; Ramon Santos and the University of the Philippines Center for Ethnomusicology for sharing their unpublished archival recordings; and Emily Merritt for coding the songs for the reliability analysis. This work was supported by the Social Sciences and Humanities Research Council (SSRHC) of Canada, grant No 90–2320-B-195–002 from the National Science Council of Taiwan, grant NHRIEX93-9218BI from the National Health Research Institute, and by an Amherst College Roland Wood Fellowship to P.E.S.

REFERENCES

Blacking, J. (1977). Some problems of theory and method in the study of musical change. *Yearbook of the International Folk Music Council*, *9*, 1–26.

Busby, G. (2006). Finding the blues: An investigation into the origins and evolution of African-American music. 1-55. (Unpublished Master's Thesis)

Cann, R. L., Stoneking, M., & Wilson, A. C. (1987). Mitochondrial DNA and human evolution. *Nature*, *325*, 31-36.

Cavalli-Sforza, L.L. (1997). Genes, peoples, and languages. *Proceedings of the National Academy of Sciences USA*, *94,* 7719–7724.

Diamond, J. (2000). Taiwan's gift to the world. *Nature*, *403*(6771), 709-710.

Dowling, W. J., & Harwood, D. L. (1986). Cultural contexts of musical experience. In W. J. Dowling & D. L. Harwood (Eds.), *Music Cognition* (pp. 225-239). Orlando: Academic Press.

Downey, J. C. (1970). Review of A. Lomax, Folk song style and culture. *Ethnomusicology*, *14*(1), 63-67.

Erickson, E. E. (1976). Tradition and evolution in song style: A reanalysis of Cantometric data. *Cross-Cultural Research*, *11*(4), 277-308.

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, *131*(2), 479-491.

Grauer, V. (2006). Echoes of our forgotten ancestors. *The World of Music*, *48*(2), 5-59.

Gray, R. D., Drummond, A. J., & Greenhill, S. J. (2009). Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science*, *323*, 479-483.

Henry, E. O. (1976). The variety of music in a North Indian village: Reassessing Cantometrics. *Ethnomusicology*, *20*(1), 49-66.

Huteson, G. (2003). Sociolinguistic survey report for the Tona and Maga dialects of the Rukai language. *SIL Electronic Survey Reports*, *12*.

Kurosawa, T. (1973). *The music of Takasago tribe in Formosa (in Japanese)*. Tokyo: Oyamakaku.

Leroi, A. M., & Swire, J. (2006). The recovery of the past. *The World of Music*, *48*(3), 43-54.

Lomax, A. (1980). Factors of musical style. In S. Diamond (Ed.), *Theory & practice: Essays presented to Gene Weltfish* (pp. 29-58). The Hague: Mouton.

Lomax, A. (Ed.). (1968). *Folk song style and culture*. Washington, DC: American Association for the Advancement of Science.

Lomax, A., & Grauer, V. (1968). The Cantometric coding book. In A. Lomax (Ed.), *Folk song style and culture* (pp. 34-74). Washington, DC: American Association for the Advancement of Science.

Maranda, E. K. (1970). Deep significance and surface significance: Is Cantometrics possible? *Semiotica*, *2*(2), 173-184.

McMahon, A. M. S., & McMahon, R. (2005). *Language classification by numbers*. Oxford: Oxford University Press.

Nachman, M. W., & Crowell, S. L. (2000). Estimate of the mutation rate per nucleotide in humans. *Genetics*, *156*(1), 297-304.

Nettl, B. (1970). Review of A. Lomax, Folk song style and culture. *American Anthropologist*, *72*(2), 438-441.

Oppenheimer, S. (2004). The “express train from Taiwan to Polynesia”: On the congruence of proxy lines of evidence. *World Archaeology*, *36*(4), 591-600.

Stock, J. P. J. (2006). Clues from our present peers? A response to Victor Grauer. *The World of Music*, *48*(2), 73-91.

Trejaut, J. A., Kivisild, T., Loo, J. H., Lee, C. L., He, C. L., Hsu, C. J., et al. (2005). Traces of archaic mitochondrial lineages persist in Austronesian-speaking Formosan populations. *PLoS Biology*, *3*(8), 1362-1372.

Tzanetakis, G., Kapur, A., Schloss, W.A., & Wright, M. (2007). Computational ethnomusicology. *Journal of Interdisciplinary Music Studies 1*(2), 1-24.

Wright, S. (1943). Isolation by distance. *Genetics, 28,* 114–138.

Chapter 5

Conclusion: Prospects for the new comparative musicology

Patrick E. Savage

This thesis has attempted to solve a number of challenges in the cross-cultural study of musical diversity. I believe that the manuscripts that comprise Chapters 2-4 complement one another in building a methodological framework that will be useful to future work integrating the qualitative research by ethnomusicologists and the quantitative research done by scientists in fields of music cognition, physical anthropology, and cultural evolution.

In Chapter 2, we presented a new cross-cultural song classification scheme. We have put much effort into developing a more theoretically grounded, consistently organized, and reliable scheme than its predecessor, Cantometrics. Hopefully, this will lead to our scheme becoming the first cross-cultural song classification scheme that is generally accepted. This scheme, in addition to serving as the basis for the musical data in Chapters 3 and 4, will also be useful for collecting empirical data to test the theoretical claims of musical universals by Brown and Jordania (in press).

In Chapter 3, we presented our new methods for quantifying musical similarity both within and between cultures. While criticisms of Cantometrics' reliability were addressed in Chapter 2, more numerous and stronger criticisms were directed at the statistical methods that were applied to the Cantometric data, particularly regarding ignorance of internal musical diversity. By adopting the Analysis of Molecular Variance (AMOVA) framework from population genetics, we not only resolve this criticism of Cantometrics, but also provide a new and important tool for people studying cultural evolution in other domains, such as language, to address the issue of intra-cultural diversity. The interesting parallel we found between musical and genetic diversity, in which the vast majority of diversity in both domains is found within cultures, also hints at the value of applying models of biological evolution to musical evolution.

In Chapter 4, we applied our new methods from Chapters 2 and 3 to directly address the relationship between musical and genetic evolution. The significant correlations we found between musical and genetic similarities among the aboriginal tribes of Taiwan support the idea that music and genes may have been co-evolving and co-migrating with genes for thousands of years. This supports Grauer's (2006) claim that music may be have the kind of deep time-depth required to study prehistoric migrations, suggesting that CantoCore, like Cantometrics, can be used to find important relationships between music and culture. However, the dissociations between correlations using structural vs. performance features and the dissociations between correlations using different sections of the mitochondrial DNA genome suggest that musical evolution, like other forms of evolution, is complex. Great care must therefore be taken in future work to account for these complexities, including finding appropriate musical samples, exploring appropriate models of musical evolution, and relating musical migration and evolution to dynamic changes in its present-day performance context.

Taiwan was chosen as a starting point for this project because of its importance in the debate about the homeland of the migration of Austronesian-speaking cultures (Diamond, 2000; Gray, Drummond, & Greenhill, 2009; HUGO Pan-Asian SNP Consortium, 2009; Oppenheimer & Richards, 2001), and because of its well-studied musical traditions (Kurosawa, 1973). Unfortunately, logistical constraints prevented us from following through on our intentions to explore other areas important to this migration, such as South China, the Philippines, Indonesia, Papua New Guinea, and the Polynesian islands. In addition to time constraints due to the fact that TR and I were only here for two years as Masters students, further complications developed regarding obtaining appropriate musical samples and accompanying ethnographic meta-data.

As a result, although I will be continuing this general line of work in my next graduate work at the Tokyo University of the Arts, I will be working with a slightly different focus using pre-existing databases such as the *Garland Encyclopedia of World Music* (Nettl, Stone, Porter, & Rice, 1998) and the *Nihon min'yo taikan [Anthology of Japanese Folk Song]* (NHK, 1944-1993). These will allow me to test both large-scale theories of global music evolution and musical diversity – such as Lomax's (1968) ideas about 10 global song-style regions evolving from 2 independent origins in Africa and Siberia – and perform my own ethnographic fieldwork to explore smaller-scale theories of musical micro-evolution and social change (Hughes, 2008). In the meantime, I will continue to explore more options for suitable musical and non-musical samples for continuing work on musical migrations in Austronesia and elsewhere. Hopefully, this work, along with the current thesis, will help stimulate a renaissance in comparative musicology that will provide fruitful material for those in ethnomusicology and music cognition that are interested in an integrated understanding of the relationship between music, biology, and culture.

REFERENCES

Brown, S., & Jordania, J. (in press). Universals in the world's music. *Psychology of Music*.

Diamond, J. (2000). Taiwan's gift to the world. *Nature*, *403*(6771), 709-710.

Grauer, V. (2006). Echoes of our forgotten ancestors. *The World of Music*, *48*(2), 5-59.

Gray, R. D., Drummond, A. J., & Greenhill, S. J. (2009). Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science, 323*, 479-483.

Hughes, D. W. (2008). *Traditional folk song in modern Japan: Sources, sentiment and society*. Kent: Global Oriental.

HUGO Pan-Asian SNP Consortium. (2009). Mapping human genetic diversity in Asia. *Science*, *326*(5959), 1541-1545.

Kurosawa, T. (1973). *The music of Takasago tribe in Formosa [in Japanese]*. Tokyo: Oyamakaku.

Lomax, A. (Ed.). (1968). *Folk song style and culture*. Washington, DC: American Association for the Advancement of Science.

Nettl, B., Stone, R., Porter, J., & Rice, T. (1998). *The Garland encyclopedia of world music (10 vols.)*. New York: Routledge.

NHK [=Nippon Hoso Kyokai] (1944-1993). *Nihon min'yo taikan [Anthology of Japanese folksongs – 13-volume CD & book set]*. Tokyo: Nippon Hoso Shuppan Kyokai.

Oppenheimer, S. J., & Richards, M. (2001). Slow boat to Melanesia? *Nature*, *410*(6825), 166-167.