CantoCore: A New Cross-Cultural Song Classification Scheme

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

Department of Psychology, Neuroscience & Behaviour, McMaster University

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" coding scheme, 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 that emphasizes its "core" structural characters rather than the more subjective characters of performance style. Using both schemes to classify the 30 songs from the Cantometrics Consensus Tape, we found that 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.

Keywords: music, song, classification, CantoCore, Cantometrics, reliability, musilinguistic continuum

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* (Stone 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:

> [C]omparative 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 and/or acoustic classification schemes, despite pleas 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 that classification schemes be universally applicable—these challenges do not *a priori* invalidate cross-cultural classification (but see Hood 1971; Blacking 1973; McLeod 1974).

Along these lines, there 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. 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. A classificatory approach based on multidimensional, musilinguistic spectra could be helpful in fields as diverse as ethnomusicology, neuroscience, and evolutionary biology for understanding connections between music and language (Darwin 1871; Feld and Fox 1994; Wallin, Merker, and Brown 2000; Patel 2008).

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 Figure 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 whenever 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.
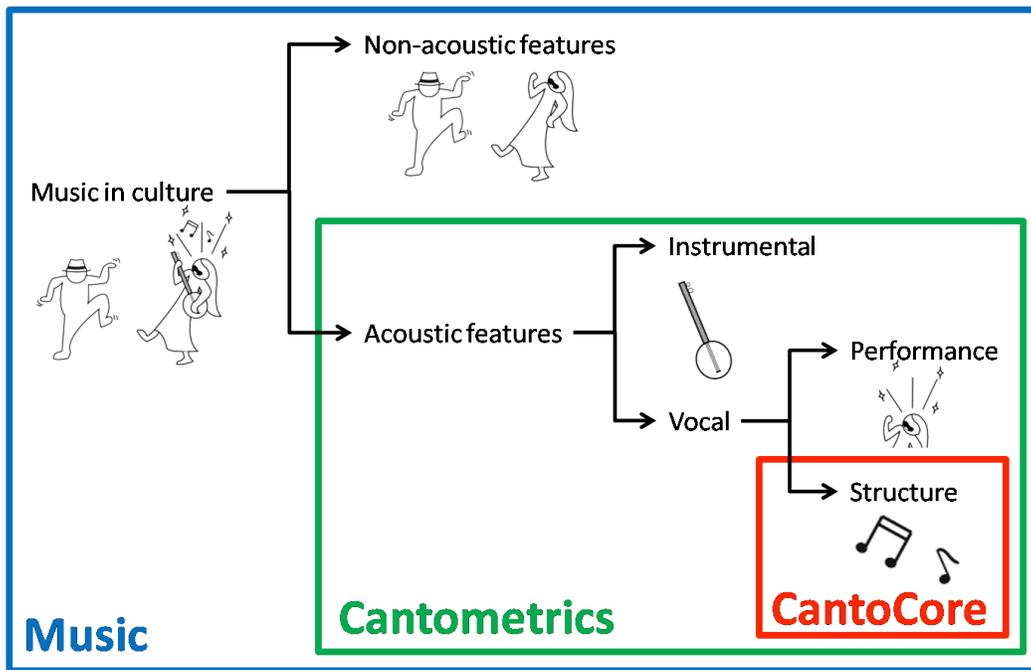
Figure 1. A comparison of the types of musical characters classified by CantoCore vs. Cantometrics. Both classification schemes rely exclusively on acoustic information rather than on non-acoustic characters. Whereas Cantometrics (green box) focuses on both the performance and structural characters of 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; the characters themselves are 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 complex combinations of notes, just as organisms are complex combinations of cells. However, as with the cells in an organism, the notes in a song
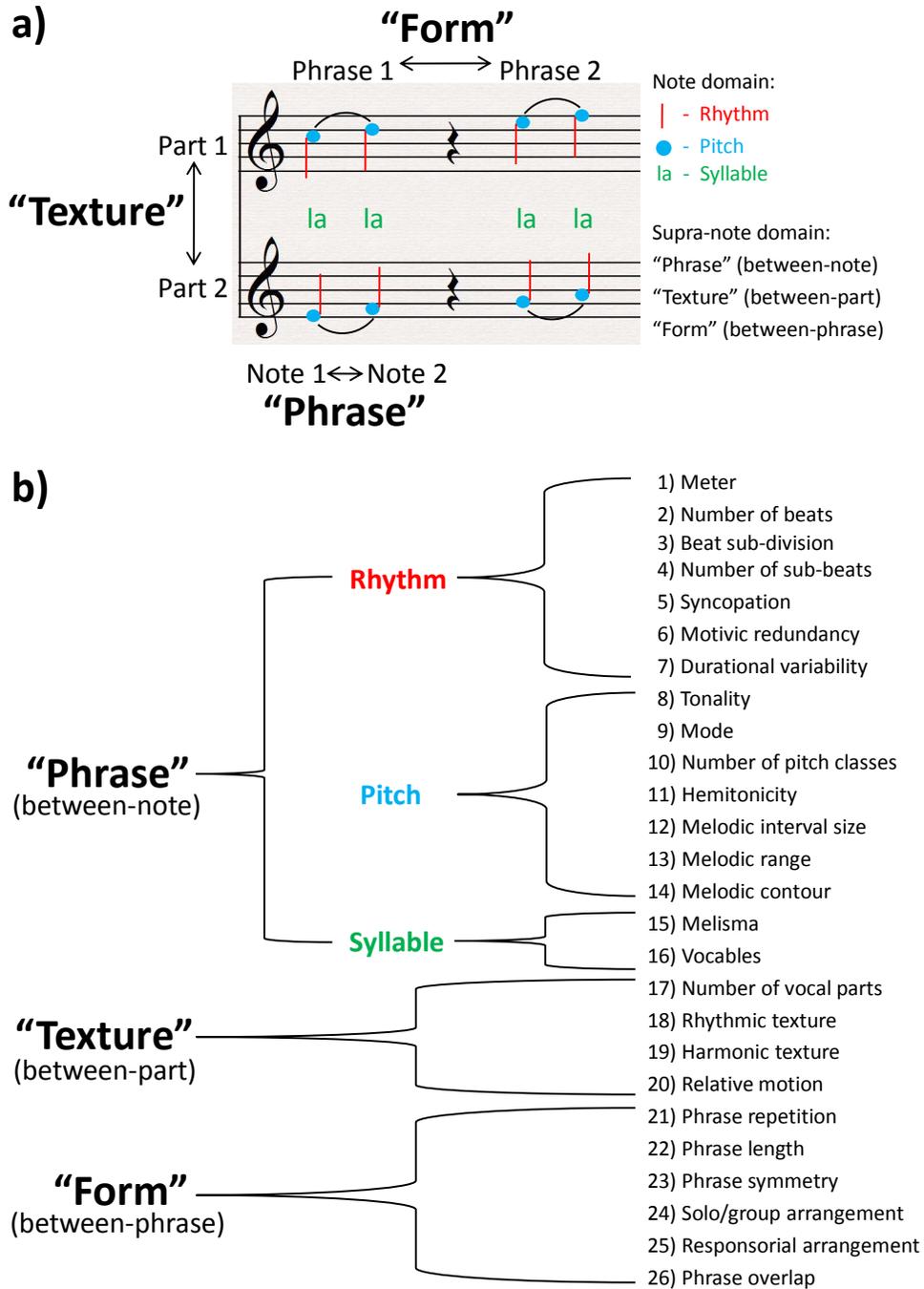
Figure 2. a) The musical hierarchy is comprised of "note" and "supra-note" domains. The three main note domains are rhythm (red), pitch (blue), and syllable (green), as represented by the sung note "la." Interactions between notes give rise to the supra-note domains of "phrase" (the between-note level), "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.

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 Figure 2a). CantoCore classifies 26 structural characters of songs (Figure 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.

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 96 character-states across the scheme. Of these, 53 are new to the scheme, as indicated by asterisks in the detailed 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. 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 Figure 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
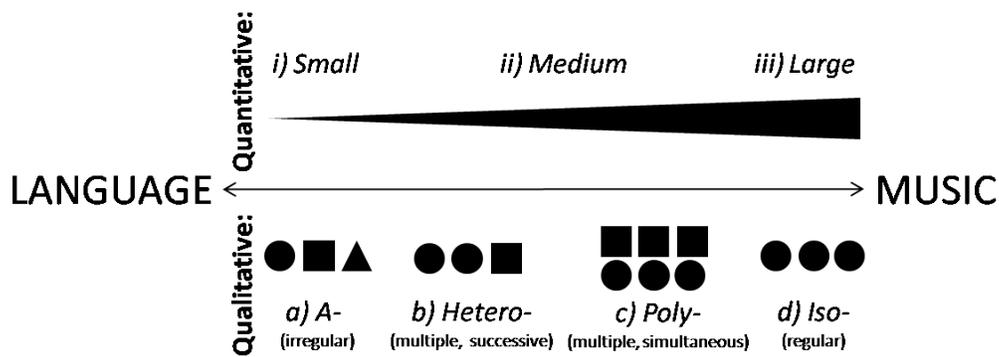
Figure 3. The character-states within each character are organized according to a "musilinguistic" spectrum spanning from language-like to music-like (no value judgment is implied). Quantitative characters (top part of the figure) are ordered in terms of increasing size from small to large using lower-case 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.

"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

Classification by ear. The goal of our classification system is to provide a tool to describe and compare songs from many cultures in terms of multiple musical features. Ideally, one would want to use an automatic acoustic-based classification system or a database of musical transcriptions/notations to allow one to quickly and objectively classify songs with a high degree of accuracy. Unfortunately, the automatic classification systems and databases that currently exist are heavily biased towards Western songs and Western theory (e.g. Schaffrath 1995; Bertin-Mahieux, Ellis, Whitman and Lamere 2011). Thus, as with the creators of Cantometrics, we have been forced to develop a relatively blunt method that can allow a coder to classify an individual song by ear across several dozen features in a short amount of time. Although we have tried to provide specific definitions and precise threshold values for all our character states, these ultimately function as rough guidelines to help the coder reach a more holistic, subjective decision regarding the appropriate classification.

The reliability of such classifications by ear is limited both by lower-level perceptual constraints and by higher-level cognitive constraints. For example, in an experiment testing interval perception among Western and Javanese musicians, Perlman and Krumhansl (1996) found great variability within both groups of musicians. Even their most accurate subjects were limited by basic perceptual constraints in their ability to reliably distinguish intervals differing by only 20 cents. On the other hand, one Javanese musician displayed "regions of confusion" as large as 180 cents in which they perceived intervals from 120–300 cents as being equivalent, presumably because they were using the *slendro* scale (which contains only one scale degree in this range) as an internal interval standard.

The greater the effects of each type of constraint, the lower the accuracy of classification by ear will be. Nevertheless, by dividing quantitative characters into only three character-states, rather than the five character-states preferred in Cantometrics, we have tried to minimize the number of grey areas where such classificatory ambiguities could occur while at the same time maintaining the sense of a continuum of musical features, rather than a "presence/absence" dichotomy.

The choice of precise threshold values is necessarily arbitrary, especially since there are no comprehensive datasets other than Cantometrics regarding the worldwide distribution of these features. Therefore, we have tried to specify values that will best capture the range of variation found throughout the world, relying mainly on Cantometrics and on our own subjective listening experiences with world musics. For example, the

use of the perfect fifth and octave as thresholds for "melodic range" (character **13**) was maintained from Cantometrics, while our choice of three and five as thresholds for the number of pitch classes in a scale (character **10**) was based on our intuitions that these would capture the most variation in scales throughout the world.

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 that 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). 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, an "n/a" character-state is included to denote something that is unclassifiable. A "?" 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 used to collect the original Cantometric data (Lomax and Grauer 1968), 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. 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 so 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 or striking characters not classifiable, 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.

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. Modifications to original Cantometrics character-states are listed using parentheses.

I) "PHRASE" (between-note)

A) Rhythm

1) **METER** (Cantometrics Line 11)
   Cyclic, hierarchical groupings of beats into bars
   (a) *A-metric:* No consistent beat (formerly *"parlando rubato – free rhythm"*)
   (b) *Hetero-metric:* There is a consistent beat, but there is no consistent hierarchical pattern among these beats (formerly divided into *"one-beat rhythm"* and *"irregular meter"*)
   (c) *Poly-metric*:* Multiple independent beats occur simultaneously (e.g., 6/8 against 3/4, multiple singers singing in different tempi) (*"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. Although it is debatable whether one can hear multiple meters simultaneously (Kolinski 1973; London 2004), it is possible for a listener to recognize the presence of two simultaneous meters/tempi and choose to attend to one or the other meter. Therefore, we have maintained this character-state, although it may not be useful in the majority of cases. 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. For example, 3/4 and 9/8 rhythms are more common in Europe than in Africa or Asia, where 12/8 and 4/4 meters, respectively, are relatively more common (Stone et al. 1998).

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.

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 that 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).

Pitch class: Notes that share the same note name (e.g., B, Db) regardless of their absolute pitch are considered as the same pitch class (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 pitch classes. Unfortunately, as a result of this compromise, there may be some cases in which separate microtones are classified as a single pitch class, while in other cases normal variation of intonation may be classified as separate pitch classes.

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.

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 q and h 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 e 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., q beat divided into e 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., q.k beat divided into e note sub-beats; includes 6/8, 9/8)

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

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

Comments: Songs in which groupings of five or more sub-beats are broken down into smaller groupings of twos and threes (e.g., 2+2+3+2+2/8 [London 1995]) should be classified as "*hetero-divisive*" songs (see **3**) rather than "*complex.*" In "swing" time, sub-divisions that approximate 2:1 should be classified as "*compound,*" while those that approximate 3:2 should be classified as "*complex.*"

5) SYNCOPATION*

The percentage of notes that are relatively prominent (loud) but in metrically unaccented positions

(i) *Little or no syncopation: <5%*

(ii) *Moderately syncopated: 5–20%*

(iii) *Highly syncopated: >20%*

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

Comments: The term "syncopation" is used here instead of Kolinski's (1973) term "contrametricity" because it is more widely understood, because it allows us to recognize a continuum of varying degrees of syncopation rather than a "commetric/contrametric" dichotomy, and because Kolinski did not offer a precise definition of contrametricity.

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 motive.

Comments: Figure 4 provides an example where 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. Duration values refer to inter-onset intervals (IOIs) as opposed to sounding durations (i.e., a quarter followed by an eighth rest and a dotted quarter both have the same IOI duration value). Dotted notes are counted as separate duration values.

Comments: Patel and Daniele (2003) present a different conception of rhythmic variability that focuses on variability between successive pairs of

notes, such that a series of all quarter notes has minimal variability and a series of alternating quarter notes and eighth notes has high variability. Our definition instead focuses on *global* variability across all notes in a song, which is easier to estimate by ear and which provides a better means of examining *differences* between music and language, whereas Patel & Daniele were explicitly trying to examine *similarities* between music and language.

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: Kolinski's (1961) scale-classification scheme did not recognize the fact that some songs have no tonic or have multiple tonics. Therefore, we have added this character to permit classification of these additional types of songs.

9) MODE*
Presence of pitch classes 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 but 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 "pitch class."
Comments: The concept of mode is complex, and the distinction between major and minor 3rd is only one of many possible angles from which to approach it (Powers et al. 2012). Nevertheless, we have chosen to focus on the major/minor distinction because it is commonly employed and relatively amenable to classification. Characters dealing with micro-tonal intonations have been avoided due to a lack of consensus on how to classify these characters.

10) NUMBER OF PITCH CLASSES*
Number of pitch classes found in the scale
(i) *Sparse scale:* <4 pitch classes
(ii) *Moderately dense scale:* 4–5 pitch classes
(iii) *Dense scale:* >5 pitch classes
(n/a) *A-/hetero-/poly-tonal:* See **(8)**
N.B. See Box 1 for the definition of a "pitch class."
Comments: The more common term "pitch class" 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 "scale degree."

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 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 greatly affect the dominant melodic contour should be ignored (e.g., changes of interval direction that last only one or two notes). Otherwise, a large

number of contours will end up being classified as "*undulating*," reducing the overall informativeness of the character.

Comments: Although most phrases in both humans and birds tend to descend in their final half (Huron 2006; Tierney, Russo and Patel 2011), three additional character-states were needed to allow for horizontal, ascending, or U-shaped contours that were 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, the current 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 and/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 feature of many musics cross-culturally but are difficult to define and code for someone 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)**. The current character instead uses words containing only vowels and/or semi-vowels as a proxy for vocables. If the coder understands the language, they may exclude lexical semi-vowel/vowel combinations (e.g., "yo-yo" in English) and/or include vocables that include consonants (e.g., Celtic mouth music).

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 (a) ("*one-part*"), including solo, unison, and doubling at the octave (i.e., "magadizing"), must be coded (n/a) for characters **(18–20)**. This does not include "multisonance" (multiple pitch classes realized simultaneously [Kolinski 1978]), whether intentional or not. "*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 **(24)** and **(18)**, 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) *Rough ("dissonant"):* 50–249 cents (includes 951–1150 cents) (e.g., 2nds/7ths)

(ii) *Smooth ("consonant"):* 250–600 cents (includes 600–950 cents) (e.g., 3rds/6ths)

(n/a) *One-part (includes 0–49 and 1150–1200 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 7th) can be transposed down one octave to create a harmonic interval of 200 cents (major 2nd). 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 rough intervals also contain smooth intervals, but not the reverse. To prevent confusion, we avoid the related terms "dissonant" and "consonant" as well as common distinctions between "consonant," "perfect" and "tritone" intervals found in Western music theory, as it is not yet established to what degree these categories are cross-culturally or experimentally valid. Plomp and Levelt (1965) developed an experimentally-based explanation for sensory consonance based on the acoustic critical

bandwidth, but there also alternative usage-oriented definitions. Although calculating critical bandwidth is impractical to do by ear, it basically corresponds to our "rough/smooth" division as well as to the traditional division in Western music theory between "consonant" 3rds/6ths and "dissonant" 2nds/7ths. It should be noted that the critical bandwidth is more complex and nuanced, varying throughout the audible range and giving different sensory consonance values for 2nds vs. 7ths.

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. The original Cantometric character contained 13 different character-states, each with a specific combination of features (e.g., *"complex strophe with little/no variation," "simple litany with high variation,"* etc.). Busby (2006) reorganized this character into three new characters— "phrase repetition," "complexity," and "amount of variation"—but we retained only the "phrase repetition" character, as we found the other two characters too difficult to reliably define and code. "*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 (Figure 4). CantoCore classifications for all 30 songs on the Cantometrics Consensus Tape are listed in Appendix A.

RELIABILITY

To compare the inter-rater reliability of each system, E.M. used both Cantometrics and CantoCore, to classify the 30 songs from the Cantometrics Consensus Tape (Figure 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 tables B1–3 for detailed results), and then averaged across all characters to compare the

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	2% (1/61 notes [2nd note of bar 9])	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) Number of pitch classes	6 pitch classes (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	13% [8/61 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	Max. of 3 new phrases (phrases 5–7) are introduced before an earlier phrase (phrase 6) is repeated	ii) Moderately repetitive
22) Phrase length	2 seconds max	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

Figure 4. Transcription of the Shona song "*Pi mcinanga*" (track 13 from the Cantometrics Consensus Tape [Lomax 1976]) and its codings on the 26 CantoCore characters. Phrases (all are two measures long) are shown using phrase marks. Syllables containing only vowels and/or semi-vowels (used as a proxy for vocables) are underlined. The actual pitches are two semitones lower than those shown in the transcription. For quantitative characters, both raw quantitative values and categorical classifications are shown. An mp3 file is available at **http://greenstone.ilam.ru.ac.za/collect/ilam/index/assoc/D11849.dir/TR174-09.mp3**

mean agreement between the two classification systems 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. For example, some amount of agreement would be expected by chance even if the coders were coding at random, but some types of disagreement (e.g., "good blend" vs. "maximal blend") are less severe than others (e.g., "maximal blend" vs. "no blend (solo)"). Furthermore, simply agreeing that a song is sung solo inflates the true agreement because a "solo" character-state is coded redundantly in six different Cantometric characters related to vocal texture and vocal blend. 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" (squared weighting) (Cohen 1968) for quantitative characters and "unweighted κ" (Cohen 1960) for qualitative

Figure 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).

characters. For both percent agreement and κ, we used only the single coding indicated as most prominent in cases of multi-coding.

For example, for Cantometrics Line 5 (tonal blend of the vocal part), the coders agreed on 21 out of the 30 codings, giving a percent agreement value of 70%. However, 10 of these cases simply represented a repetition of an agreement that there was only one vocal part, a character-state that had already been coded in Line 4. When we limit this character's analysis to only the 18 characters where both coders had already agreed in Line 4 that there were in fact multiple voices to base an estimation of tonal blend upon, the percent agreement value would be 61%. However, this does not account for the amount of chance agreement we would expect given each coder's baseline propensity for choosing each character-state (the "joint marginal probability"), which in this case is 31%. Cohen's Kappa effectively calculates the proportion of agreement after subtracting out this chance agreement as follows: $\kappa = (0.61–0.31) / (1–0.31) = 0.43$. However, this still does not account for the degree of partial agreement in cases where the coders did not exactly agree. For example, "4) good blend" is three times closer to "5) maximal blend" than is "2) no blend [solo]." With squared weighting, disagreement between 5 and 2 is weighted as 3^2, i.e., nine times as severe as disagreement between 4 and 5. Once we also incorporate information about degree of weighted

partial agreement (including both the observed partial agreement and the joint marginal probability of this agreement), we arrive at a final weighted κ value of 0.49.

The results for κ are shown in Figure 6. 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 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).

There is some debate about how to interpret kappa-statistics, as Landis & Koch's criteria, although useful, are self-admittedly arbitrary. Some authors have proposed further additions to κ, such as using the maximum attainable κ given the coders' pre-existing marginal probabilities or using a minimum acceptable threshold value for κ (e.g., 0.4 for clinical uses) rather than zero (Sim & Wright 2005). However, it should be noted that Cohen
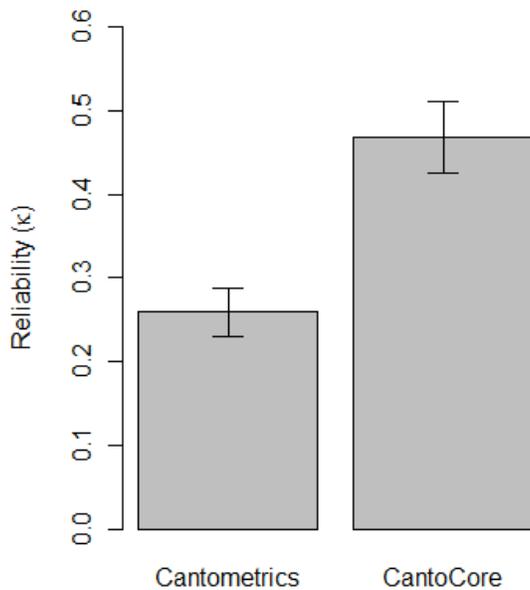
Figure 6. Mean reliability for all 37 classification characters in Cantometrics and all 26 characters in CantoCore. Error bars represent the standard error of the mean. CantoCore is significantly more reliable than Cantometrics (p=0.0001).

(1960) originally advised against giving much weight to maximum attainable κ, as "disagreement which is forced by marginal disagreement has the same negative consequences as that not so forced—in short, it is disagreement," and that a minimum acceptable threshold value of 0.4 is equally as arbitrary as Landis & Koch's criteria.

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) (Tables B1–2). Therefore, it may still be useful to supplement CantoCore's structural characters with the performance characters from Cantometrics. Cantometrics' instrumentation characters (Table B3), however, were almost three times less reliable than its structural and performance characters (instrumentation: mean κ = 0.11). This is consistent with our prediction that songs are more amenable than instrumental music to reliable cross-cultural classification. We could not reject the null hypothesis that agreement on instrumental characters was simply due to chance at the standard significance threshold of p=0.05, although the obtained value of p=0.07 is so close to this threshold that it may well have been significant given a larger sample size.

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 be a stronger reflection on our coder and/or our training procedure 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. One possibility to improve reliability for both systems in the future would be to have multiple independent coders classify songs and construct a consensus coding based on their combined agreement.

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 2012) provide new methods that computational ethnomusicologists can use 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 (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, including the forces of musical change and stasis both geographically and historically. When attempts have been made in this direction (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 2011).

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

REFERENCES

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

Agawu, Kofi. 2010. "Against Ethnotheory." Paper presented at the First International Conference on Analytic Approaches to World Music, Amherst, Massachusetts, February 19–21.

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.

Bertin-Mahieux, Thierry, Dan P.W. Ellis, Brian Whitman, and Paul Lamere. 2011. "The Million Song Dataset." *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011)*: 591–596.

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, MA: MIT Press.

Brown, Steven, and Joseph Jordania. 2011. "Universals in the World's Musics." *Psychology of Music*. doi:10.1177/0305735611425896.

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.

_____. 1978. "The Structure of Music: Diversification Versus Constraint." *Ethnomusicology* 22.2: 229–244.

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

Lomax, Alan, and Norman Berkowitz. 1972. "The Evolutionary Taxonomy of Culture." *Science* 177.4045: 228–239.

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.

London, Justin. 1995. "Some Examples of Complex Meters and their Implications for Models of Metric Perception." *Music Perception* 13.1: 59–77.

———. 2004. *Hearing in Time: Psychological Aspects of Musical Meter*. Oxford: Oxford University Press.

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.

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.

Perlman, Marc, and Carol L. Krumhansl. 1996. "An Experimental Study of Internal Interval Standards in Javanese and Western Musicians." *Music Perception* 14.2: 95–116.

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.

Powers, Harold S., Frans Wiering, James Porter, James Cowdery, Richard Widdess, Ruth Davis, Marc Perlman, Stephen Jones, and Allan Marett. 2012. "Mode." *Grove Music Online. Oxford Music Online.* http://www.oxfordmusiconline.com/subscriber/article/grove/music/43718#S43718 (accessed March 1, 2012).

Rzeszutek, Tom, Patrick E. Savage, and Steven Brown. 2012. "The Structure of Cross-Cultural Musical Diversity." *Proceedings of the Royal Society B: Biological Sciences* 279: 1606-1612.

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

Schaffrath, Helmut. 1995. *The Essen Folksong Database in Kern Format*, ed. D. Huron. Menlo Park, CA: Center for Computer Assisted Research in the Humanities.

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

Sim, Julius, and Chris C. Wright. 2005. "The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements." *Physical Therapy* 85.3: 257–268.

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

Stone, Ruth, Bruno Nettl, James Porter, and Tim Rice, eds. 1998. *The Garland Encyclopedia of World Music (10 vols., 9 CDs)*. New York: Routledge.

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

Tierney, Adam T., Frank A. Russo, and Aniruddh D. Patel. 2011. "The Motor Origins of Human and Avian Song Structure." *Proceedings of the National Academy of Sciences* 108.37: 15510–15515.

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, MA: 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. Recordings of these songs are available on Tape VII of Lomax (1976), and are scheduled to be re-released digitally by the Association for Cultural Equity at http://research.culturalequity.org/cls.jsp.

	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	**ce**	ii	ii	ii	iii	a**c**ef	i	ii	iii	b**c**	ii	ab**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	ae	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	ae	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	ab**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

Cantometrics Consensus Tape track number (row labels) / **CantoCore character number** (column headers)

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
Number 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 pitch classes	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