Chapter 13

Beyond Harmonicity: Toward a Vocal-Melodic Theory of the Origin of Musical Scales

Elizabeth Phillips and Steven Brown

For much of Western intellectual history since the time of the ancient Greeks, music has been considered as a science. During the medieval period, music was one of four liberal arts that made up the "quadrivium" of arithmetic, geometry, astronomy, and music (Kristeller 1951). This perspective of music as a science was based on the observation, credited to Pythagoras, that harmonic intervals could be mathematically described as simple proportions of string lengths. Music, as a system of perfect harmonic proportions, was seen as rooted in nature itself, not just in the human mind. In the Enlightenment period, however, music's standing changed radically from "science" to "art." In particular, it was considered as one of the "fine arts" defined by Batteux in 1746. This reframing involved, among other things, a switching of emphasis away from musical instruments and towards the voice (Berger 2000). Vocal music was increasingly seen as an ideal medium for expressing emotion (Rousseau 1781/1998), or for portraying "a picture of the human heart" (Batteux 1746/2015). This idea was manifested in the rise of opera and art songs, as well as by intellectual inquiry into the connection between music and speech, both vocally and emotionally. Whereas the ancient Greeks' philosophy of music was dominated by harmony, acoustics, and mathematical ratios, later views gave precedence to melody, performance, and emotion, especially as produced by the voice (Thomas 1995; Brown 2007).

The fields of psychoacoustics and music cognition, which began to flourish through the work of perceptual scientists and physicists in the nineteenth century, are currently grounded in an ancient-Greek-inspired "harmonicity" view of music. This grounding can be seen quite clearly in how musical scales have been studied and discussed for several hundred years. For example, the accounts of musical scales developed by Rameau (1722) and Helmholtz (1885) both argue that the pitches comprising diatonic scales are derived from ratios within the natural harmonic series. This is a theory-driven perspective of scales in that it defines a series of pitch classes *a priori* that can exist independent of particular melodies. From this perspective, scales are the set of possible notes, selected according to perfectly proportioned harmonic intervals, from which melodies and harmonies can be crafted. Music that fails to abide by these perfect ratios is considered poorly tuned, chromatic, or atonal. These *a priori* scales have served as the blueprint for most Western instrumental tuning systems. Likewise, music cognition research has consistently prioritized the music of fixed-pitch instruments, often in Western classical or popular forms. The Western major and minor modes, for example, are very nearly the exclusive focus of study in music cognition research. As such, the "harmonicity" view of scales, featuring an emphasis on mathematics and engineering, has been the default perspective for many music researchers, even those examining the parallels between music and the voice (cf. Patel 2008).

The "harmonicity" view of musical scales can be contrasted with an alternative perspective that prioritizes melody and physiology above harmony and physics. In such an account, scales are *a posteriori* descriptions of the way that people actually produce music, most especially melodies. Variable-pitch instruments have flexible tuning systems, and the voice in particular has no *a priori* tuning. In most acoustic music production, pitches are imprecise and consist of a range of frequencies (Mason 1960; Rakowski 1990; Vurma and Ross 2006) resulting from naturally imperfect physiological mechanisms, instrument engineering, or a combination of the two. Scales thus consist of the pitch classes that emerge as stable anchor points throughout the variable contours of melodic motion. These pitch classes are not perfect singular frequencies, but instead categorical distributions of frequencies, which need only be perceptually distinguishable from their nearest neighbors (Pfordresher and Brown 2017) while retaining membership in the melodic group (Ahlbäck 2004; Schmuckler 2009). This view of scales gives priority to vocal music, where scales can be intimately connected with the physiology of both speech and the vocal expression of emotion. As a descriptive model, this view is effective at accounting not only for Western diatonic scales, but scales from other tuning systems and even the scales of non-literate musical traditions. In essence, this practical model looks beyond the mathematical precision of perfect proportions, and instead confronts the highly imperfect nature of both musical production and cognition, which are characterized by imprecision, drift, categorical perception, and schematic memory.

Historically speaking, the "harmonicity" model has been the dominant view of scales for centuries. There is a wealth of literature claiming that scales are derived from the intervals of the harmonic series, although there are also numerous studies documenting the theory's weaknesses (Kreitler and Kreitler 1972; Parncutt and Hair 2018). We will refer to this body of literature as Harmonicity theory in this chapter, and will contrast it with a vocal-melodic hypothesis called Interval Spacing theory (Pfordresher and Brown 2017). Figure 1 provides a brief summary of the opposing tenets of each theory. In the following sections, we will compare the approaches, conclusions, and biological rationales of one landmark study from each theory as vehicles for discussing the supporting evidence thereof. We will then present the cognitive and evolutionary implications of the Interval Spacing theory relative to Harmonicity theory.

Harmonicity theory	Interval Spacing theory
Music as a science	Music as an art
Fixed-pitch instruments	The voice
Mathematical perception	Emotional communication
Physics	Physiology
Harmony	Melody
Pitch classes	Interval classes
Prescriptive: a priori tuning	Descriptive: a posteriori tuning
Top-down mechanism	Bottom-up mechanism
Divisional	Combinatoric
Harmonic compression	Interval island model
Perfect harmonic ratios	Vocal imprecision
No concept of precision	No concept of accuracy

Figure 1: A summary of Harmonicity theory and Interval Spacing theory, presented as dichotomies between the viewpoints, approaches, and rationales of each theory regarding musical scale structure.

Notable Empirical Studies: Approaches and Conclusions

In one landmark Harmonicity study, Gill and Purves (2009) defined a relatively exhaustive set of synthetic pentatonic and heptatonic scales comprising combinations of justly-tuned frequency ratios. They then calculated the mean similarity of every pairwise interval within each scale to the harmonic series, resulting in a "harmonicity score" for each scale. The authors claimed that most important scales from around the world ranked among the most harmonic scales they produced during this process, although only 9 of the 40 most harmonic 7-tone scales were identifiable as common Western, Arabic, or Indian classical scales. Notably, the most harmonic of these scales was Phrygian, a mode reflecting Pythagorean just intonation, rather than the major or minor scales that are dominant in Western music.

It should be obvious that if one *only* examines theoretical scales that have been synthesized entirely from intervals derived from the harmonic series, then one can *only* conclude that harmonicity is responsible for scale structure. These are neither surprising nor falsifiable results. The results of Gill and Purves and many others in the field (for example, Cartwright et al. 2002; Honingh and Bod 2011; Durfee and Colton 2015) could thus be marred by confirmation bias. The most important thing to acknowledge about Harmonicity theory is that it is designed to perfectly describe synthetic justly-tuned scales, *but no others*. This is a serious limitation, since even Western theoretical scales are no longer justly-tuned. Twelve-tone equal temperament (12-ET) is by far the dominant tuning system in Western music, and the intervals it contains are not small-integer frequency ratios (Martin 1962). Rather, 12-ET divides an octave (stretched, not justly tuned) into twelve equal steps of 100 cents apiece, thus defining notes as a range of frequencies, and intervals as the distance between notes. Overall, Harmonicity theory provides a good foundation for justly-intoned scales, but its application outside such scales may be quite limited.

In contrast to this purely theoretical approach, Pfordresher and Brown (2017) analyzed the structure of music actually produced by singers, in this case both trained and untrained Western singers. They analyzed the tuning accuracy and precision of the sung intervals of the familiar song "Happy Birthday" alongside accuracy judgments from average listeners. While trained singers produced the melody with slightly more-accurate 12-ET tuning than untrained singers, the accuracy of all singers was significantly lower than that of instruments that are actually engineered for 12-ET. The intervals between melodic notes were quite imprecise as well, to the point that the distributions of each interval class were at least 200 cents wide on each side of the peak, and strongly overlapped with the neighboring interval class distributions. However, despite this quantitatively low melodic accuracy and precision, the listeners consistently reported that the singing sounded good and correct.

Pfordresher and Brown's experiment showed that scales can be related to the mechanisms of pitch and interval production. The constraints of the human vocal mechanism, being relatively imprecise compared to the tuning properties of fixed-pitch instruments, shape the structure of vocal scales. This observation led to the proposal of the Interval Spacing model of scale generation, which argues that the imprecision intrinsic to vocal interval production imposes a lower boundary on the size of scale intervals, so that adjacent pitch classes can be distinguishable from one another in both production and perception. Compared to the prescriptive nature of Harmonicity theory, which serves to tune instruments *a priori* according to perfect harmonic ratios, the Interval Spacing model is a descriptive theory based on *a posteriori* analyses of the actual pitch patterns produced by singers in real-world contexts. The observational nature of this theory gives rise to some complications, most notably a potential chicken-and-egg problem regarding where the pitch classes of scales originate from. We will discuss some of these complications and possible solutions at the end of the chapter.

Biological Rationales for Scale Theories

In this section, we will compare Harmonicity theory and Interval Spacing theory with regard to their biological rationales and practical applicability. Although both mechanistic explanations centralize the voice, the Harmonicity narrative, while parsimonious, lacks empirical evidence and focuses solely on interval *perception*. Interval Spacing theory is more complex since it consolidates biological evidence about pitch production, interval production, perception, memory, and transmission. The essential difference between the mechanisms underlying the Harmonicity and Interval Spacing theories is that the former prioritizes harmonic intervals, while the latter prioritizes melodic intervals. Harmonicity theory proposes that scales are derived, via psychoacoustic perception, by compressing intervals from the harmonic series. Rameau, one of the founders of Harmonicity theory, asserted this viewpoint when he claimed that "a knowledge of harmony is sufficient for a complete understanding of all of the properties of music" (1722/1971, 3). By contrast, Interval Spacing theory proposes that scales are essentially melodic schema, and that they therefore emerge from reiterated melodic intervals.

Intervals as Harmonic Ratios vs. Intervals as Categorical Distances

Gill and Purves (2009) proposed that the humans have a perceptual preference for scales with greater overall similarity to the harmonic series, due to our experience with speech sounds. Since speech is pitched, every time we listen to speech we hear the harmonic series, which can then serve as the basis for musical scales. Kreitler and Kreitler (1972) refer to this idea as "overheard overtones." Since conspecific vocalizations are arguably the most evolutionarily relevant auditory stimuli in the environment, our brains may have evolved a perceptual preference for the harmonic intervals contained in speech, and thereby prefer scales that contain as many of these intervals as possible. However, this hypothesis raises an important question. If pitched sounds containing features of the harmonic series naturally lent themselves to the evolution of musical tonality, then *why isn't speech itself tonal*? Why would speech merely be the source of tonality but not be the manifestation of it? It seems that speech (especially in non-tonal languages) and tonal music (especially of pre-tuned intervallic instruments) exist on a continuum, with tonal languages, poetry, chant, song, and non-intervallic music lying in between these two extremes (Feld and Fox 1994; Savage et al. 2012; Brown 2017).

Even when considering psychoacoustics alone, Harmonicity theory faces the critique that the individual harmonic components of pitched sounds may not be independently perceptible by humans. Our ears do not possess the frequency discrimination capabilities necessary for overt harmonic perception. Rather, harmonic fusion occurs early in pitch processing, and the unique harmonic profile of a note contributes to our pitch perception only covertly via the percepts of pitch and timbre (Pantev et al. 1996). Listeners do not typically perceive individual frequency ratios from the harmonic series, as harmonics are perceived categorically (Parncutt and Hair 2018). The existing evidence for individual harmonic perception suggests that, at best, the lowest harmonics of a pitched tone are salient to varying degrees. However, Western listeners (perhaps from a lifelong exposure to 12-ET) demonstrate a strong preference *against* justly-tuned versions of even the lowest and most salient harmonic (the octave) as it is (Burns 1999). Moreover, constructing the entire diatonic scale from the harmonic series would require that the ear hear and collapse pitches from across *several octaves* of harmonics. Evidence suggests we are incapable of doing that.

The claim that humans innately prefer harmonic intervals is also largely unsupported by some commonly observed properties of scale structure. Harmonicity theory implies that any scale comprised of intervals from the harmonic series would be preferred, no matter how many harmonics are used or how far the harmonics are from the fundamental. Indeed, the more harmonic intervals, the better. Yet, world scales tend not to use more than 12 primary divisions of the octave or include scale intervals smaller than 100 cents. Most scales only consist of 5-7 scale degrees (Savage et al. 2015). This apparent upper limit on the number of scale degrees, combined with a lower limit on the distance between scale degrees, is not accounted for by Harmonicity theory, which could theoretically continue to divide the octave *ad infinitum*.

If the most preferred intervals for scales are those with small-integer frequency ratios, then why do melodies (and small scales based on them) contain so few octaves, fifths, and fourths? The majority of intervals used in melodies and the majority of intervals between adjacent scale degrees are small and step-like (Ortmann 1926; Dowling 1978; Vos and Troost 1989; Li and Huron 2006). Harmonicity theory would predict that for a scale consisting of only two notes, the preferred interval between them would be an octave (the smallest integer ratio for two notes); for a scale of three notes, the intervals would be an octave, fifth, and fourth (for example, the distances between C, G, and C) (Kreitler and Kreitler 1972). This prediction is not validated by any world scale that we are aware of. Rather, non-literate small scales—including, for example, those produced by birds, children, and the musicians of predominantly oral traditions—are primarily comprised of the same small and inharmonic intervals that dominate melodic motion, namely seconds and thirds (Sachs 1943; Savage et al. 2015).

In contrast to Harmonicity theory, Interval Spacing theory is predicated on how the production and perception of melodic intervals influence scale structure. Faced with the fact that intervals could be produced imprecisely and still be perceived as accurate, Pfordresher and Brown (2017) concluded that intervals cannot be defined strictly as the punctate, exact frequency-ratios that mathematical acoustics idealizes, nor can scales be viewed as ladders with those ratios as precise rungs. They instead visualize scales as a series of stepping stones or islands, where each interval has a width and distance relative to every other. Scales, in general, are defined by the number of pitch classes they contain and the relative distances between them.

This view argues that the distance between pitch classes is fundamentally codependent on the precision (or width) of both the pitch classes themselves and the intervallic movements between them during sung melodies. Combining this proposal with observed sensorimotor and memory constraints offers a biologically-relevant explanation for the properties of melodic motion, and thereby of scale structure itself. If notes and intervals have a minimum width of roughly 100 cents (+/– 50 cents) due to vocal imprecision, and if leaps larger than half an octave are used sparingly in both vocal music and speech due to production effort, then musical scales can be viewed as *a trade-off between distinguishability and reproducibility*. Scale degrees must be far enough apart that their distributions do not overlap to the point of being indistinguishable, but not so far that the interval is difficult or undesirable to produce. The lower limit on the size of intervals imposed by this trade-off automatically entails an upper limit on the number of scale degrees that can fit within an octave. Pfordresher and Brown (2017) argued that scales containing 5-7 notes per octave—in other words, scales having adjacent pitch spacings of roughly 200 cents—were exactly the outcome predicted from vocal imprecision during interval production. Harmonicity theory is unable to account for the size of scales and their intervallic spacings.

Several studies have provided evidence for the claim that people both perceive and produce pitches and intervals as categorical islands, not as single frequency ratios. Although natural tuning precision depends on multiple factors, including the instrument itself and the musician's level of training (Mason 1960; Ward 1970; Geringer and Sogin 1988), musical notes can best be described as distributions of sound frequencies. Likewise, the perception of pitch precision and accuracy depends on many factors, including the listener's musical experience. Studies have repeatedly shown that tuning, or fine-grained pitch precision, is subject to categorical perception (Siegel and Sopo 1975; Siegel and Siegel 1977a; Siegel and Siegel 1977b; Zatorre and Halpern 1979; Loosen 1994, 1995; Morrison 2000; Vurma and Ross 2006).

Vocal music in particular has no *a priori* tuning and is full of complex melodic and intonational gestures, such as vibrato, melisma, and long-range shift (Mauch et al. 2014). Notes produced by the voice are especially imprecise compared to those of fixed-pitch instruments (Pfordresher et al. 2010; Pfordresher and Brown 2017). In parallel, our perception of vocal intonation seems particularly categorical, a phenomenon known as "vocal generosity" (Hutchins, Roquet, and Peretz 2012). There is even evidence that vocal melodies with unusually precise intonation (that is, lacking vibrato or other microscopic forms of pitch variation) are perceived as incorrect, out of tune, or unnatural (Geringer, MacLeod, Madsen and Napoles 2015). Harmonicity theory views this pitch imprecision as an error surrounding a perfect mathematical ratio. Interval Spacing theory, on the other hand, views it as a *characteristic* feature of music (and especially of vocal music), one that drives the structure of melodic intervals and therefore the structure of scales.

Sensory Consonance vs. Schematic Memory Construction

Although intonation studies have demonstrated categorical perception of pitch, some researchers use the percept of sensory dissonance that can arise from inharmonic simultaneous intervals to argue that humans are nonetheless extremely sensitive to frequency imprecision (Terhardt 1984; Schellenberg and Trehub 1996). Recall again that the vast majority of Western music today is composed in 12-ET, and that nearly every musical interval we hear consists of imprecise frequency ratios. This normalized intonation inaccuracy does not appear to drastically reduce the perceived quality of the music (Warren and Curtis 2015). Moreover, sensory dissonance has only been shown to arise from simultaneous (harmonic) intervals (cf. Kameoka and Kuriyagawa 1969; Hutchinson and Knopoff 1978; Huron 1994), not from sequential (melodic) intervals (Schellenberg and Trainor 1996). This distinction likely exists because sensory dissonance mostly relies on harmonic beating, which occurs when harmonics are misaligned. However, beating is not even the best indicator of mistuning at times (Hall and Hess 1984) because harmonics are usually perceived in categorical groups (Terhardt 1987; Parncutt 1989). Even when sensory dissonance *does* indicate mistuning, it seems to be a learned rather than psychoacoustic association (Huron 2004; Mcdermott et al. 2016), because the importance of sensory dissonance varies widely across cultures (Ellis 1885; Ellis 1965; Cazden 1945, 1980; Mcdermott et al. 2016). Therefore, a bias against sensory dissonance and towards harmonic intervals is unlikely to be the innate mechanism driving human scale evolution, as Harmonicity theory suggests (Burns and Ward 1978). In fact, it generally seems unlikely that simultaneous intervals are better candidates for building scales than sequential intervals. When all of the white keys in one octave on the piano are played at once, one does not perceive the major scale. At best, one perceives a tone cluster. Some sequence of these notes seems necessary to form the impression of a scale.

Interval Spacing theory aligns well with a proposed cognitive model that views scales as heuristic schemata reflecting the sets of sequential intervals we typically hear in melodies (Dowling 1978; Dowling and Bartlett 1981; Cross, Howell, and West 1983; Jordan and Shepard 1987). According to this model, when we listen to a melody, we bin the continuous domain of frequencies into categorical pitch classes (Burns and Ward 1978; Burns 1999), each of which contains the entire frequency distribution of what we perceive as a single note. We store information about the notes we heard in both the time and pitch domains such that—within veridical memory constraints—we retain both the temporal sequence and the categorical identity of the notes we heard (Deutsch 1969). With enough iterations of the same or similar melodies, veridical knowledge gives way to heuristic knowledge to optimize information processing (Terhardt 1991).

In this process, the melody is collapsed or abstracted across the time domain such that a schema is formed with only the set of notes, rather than their sequence (Bharucha and Todd 1989). Since absolute-pitch memory is rare, or at best latent in most listeners, these schemata generally represent the relative spacing of the note categories, consisting of their individual widths and the distances between them (Dowling 2002). These schemata are the simplest representations of a musical scale. If the scale is anchored to a particularly salient note, it may represent a mode. If different notes of the mode are given different harmonic or melodic importances, it may represent a key or tonal hierarchy (Dowling 1978). However, these complications need not be assumed in cultures where their presence or relevance is uncertain (Sachs 1943). A simple descriptive model of the scale suffices, and offers a flexible explanation of melodic systems across a variety of cultures. Due to its reliance on vocal sensorimotor mechanisms, the model is especially relevant when considering the evolutionary intersections between speech and language processing (Patel 2008) and a proposed joint "musilanguage" precursor (Brown 2017).

Figure 2 provides a comparative analysis of Harmonicity theory and Interval Spacing theory with regard to the origin of scales and melodies. According to psychoacoustic theories like Harmonicity theory, the pitch classes of musical scales are established *a priori* so as to optimize the overall harmonicity of the intervallic ratios. Instruments are tuned to these ratios, and melodies are then generated based on these prior tunings. According to the Interval Spacing theory, scales are an abstraction of the way people single melodies. Physiological factors related to the imprecision of vocal pitch production provide constraints on the intervallic properties of melodies. Bottom-up combinatorial mechanisms, based on spacing principles, optimize the interval combinations that generate scales.

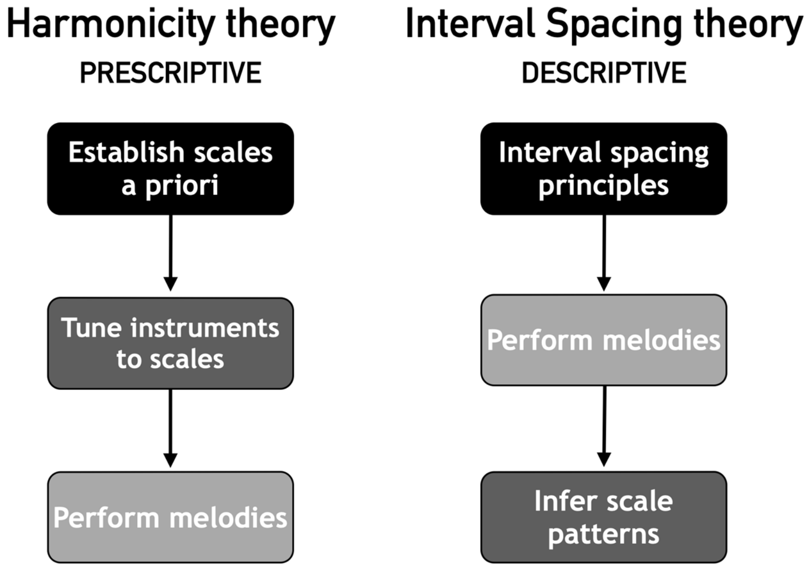

Figure 2: A summary of the relationship between scales and melodies according to Harmonicity theory and Interval Spacing theory.

Cognitive and Evolutionary Implications of Interval Spacing

The proposed biological and evolutionary mechanisms of Harmonicity theory require perfectly precise *and* accurate pitch and interval production, as well as a highly sensitive pitch-perception mechanism. The empirical evidence regarding music production and perception has shown that acoustic music usually fails to meet those requirements, and that listeners may not be sensitive to extremely fine-grained pitch differences in most real-world scenarios. Although Harmonicity theory succeeds at describing theory-based Western instrumental scales, its applicability to other world scales, especially in non-literate musical cultures that lack aesthetic treatises on scale theories, is uncertain at best. We do not believe anyone would argue that Western instrumental art music is the most prototypical, let alone evolutionarily ancient, musical tradition. As such, we would like to develop an alternative theory of scale structure that accounts for the empirical evidence regarding the physiology and cognition of musical scales, and that draws from a wide variety of global musical traditions. In particular, we would like to elaborate on the properties of the Interval Spacing theory as a descriptive, vocal, and melodic theory of musical scales.

Why a Descriptive Theory?

Interval Spacing theory, by operating under an *a posteriori* conception of scales, aims to expand the scope of scale research from the Western theoretical realm to the real world of global human music, at least as it has been recorded to date. As a descriptive theory, it has greater flexibility for cross-cultural analysis. It encourages researchers to investigate the properties of global scales that a prescriptive theory like Harmonicity cannot. Examples include small scales that do not span an octave, scales that do not exhibit octave equivalence, scales that mutate depending on melodic motion, and other scales that do not fit within a Western theoretical framework. These examples need to be studied as legitimate examples of scales, rather than as deviants.

Most importantly, a wide lens on scale structure enables the development of an explanatory account that is grounded in the fundamental commonalities and constraints of human physiology. This has the potential to offer a rich *explanatory* account of world music. The physiological mechanisms of music production that we all share could restrict us to a few basic, relatively-universal building blocks for song. Just as the theory of generative phonology in linguistics argues that the limited palette of phonemes producible by the human voice gave rise to over 7,000 languages, similar combinatorial mechanisms could be responsible for the near infinite variety of musical traditions around the world (Kenstowicz and Kisseberth 2014). Finding these universal building blocks of music requires that we expand our focus beyond the Western tradition.

Interval Spacing theory not only includes but explicitly centralizes descriptions of musical traditions that have been largely ignored outside of ethnomusicology. In doing so, it encourages a diversification of corpus studies in empirical musicology and an expansion beyond the Haydn-Bach-Mozart-Beethoven canon that has nearly single-handedly informed scale theories for several centuries. Incorporating non-Western traditions, however, often requires a descriptive theory of musical scales. Many indigenous musical traditions are non-theory-driven (at least in explicit writing) and are based in veridical oral transmission. Scales exist only in so far as they can be derived *a posteriori* from melodies (Burns 1999; Ellis 1965). Unfortunately, such scales, as with phonemes, can only be derived by ear, thus introducing the cognitive and cultural biases of the listener. Even when recorded melodies from such cultures are analyzed using computational methods, such analyses can be difficult to execute and interpret. Descriptive scale theories, while oftentimes the *only* valid scale theory, are just as detailed and exploratory as the observational methods used to generate them.

In our studies investigating Interval Spacing theory, we have attempted to strike a balance between computational methods and informed listening. We use the YIN pitch-tracking algorithm (de Cheveigné 2002) to represent the exact pitch and time dimensions of the melody. This results in a melograph (first introduced by Schroeder in 1968) that can be used to examine the patterns, features, and distributions of the melodic notes, and how they may reduce to the categorical pitch classes that comprise the scale. Figure 3 presents an example of this using the song "Twinkle, Twinkle, Little Star." Melography might even be considered as score-study, or notation musicology, although it uses a form of "global notation," rather than Western staff notation (Killick 2020). Previous computational studies have used recordings and automatic pitch-tracking algorithms, but have derived scales from pitch-class histograms, rather than melographs (Panteli et al. 2017; Chiba et al. 2019; Kuroyanagi et al. 2019; Mcbride and Tlusty 2019). Unfortunately, this method has always assumed octave equivalence. That is, the "pitch classes" in these studies are more accurately described as pitch chroma in which all of the melodic pitches are collapsed into the range of a single octave. Even if these models can effectively represent Western melodies, there is no evidential basis for using pitch chroma, rather than pitch classes, in cross-cultural research, where octave equivalence cannot be assumed.

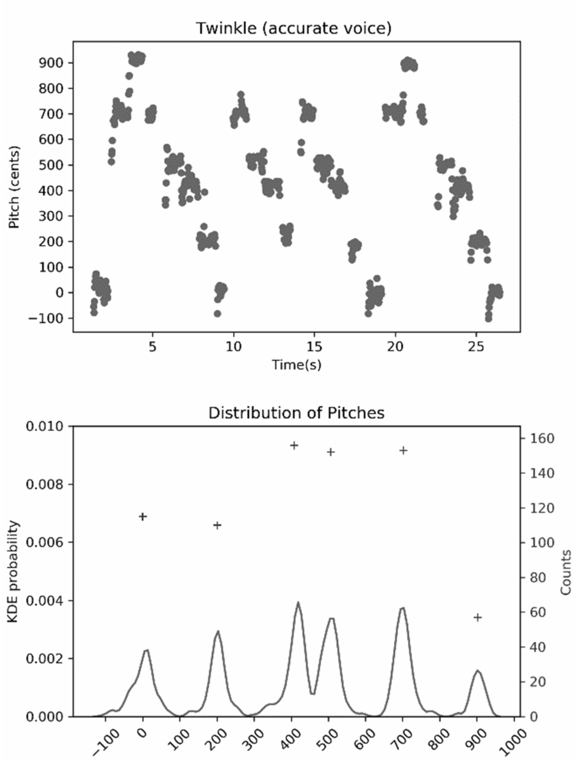

Figure 3: Top: the melograph of a recording of a trained singer performing "Twinkle, Twinkle, Little Star." Bottom: the associated pitch histogram of the same recording. Unlike in many previous studies, the pitch histogram is both smoothed and unwrapped (extending lower than the tonic, rather than being contained within a single octave).

These studies have also seemingly failed to investigate data for patterns other than those posited by Harmonicity theory, despite having access to rich quantitative data on the exact frequencies in each recording, give or take some pitch-tracking errors. For example, if a pitch histogram demonstrates a pitch chroma peak roughly 700 cents from the listener's chosen tonic, then researchers claim this as evidence for a human preference for small-integer frequency ratios. However, such "perfect" pitch distributions are typically no less than 100 cents wide (often times much greater) and are often considerably displaced from a round 100-cent interval from the tonic, demonstrating clear imprecision and inaccuracy relative to harmonicity's "perfection." Ignoring this multidimensionality reflects a deep confirmation bias in the field. To counter this bias, we are using Interval Spacing theory to generate new hypotheses about the pitch patterns in music and the justifications for investigating them.

Why a Melodic Theory?

Interval Spacing theory is descriptive in that it derives scales from existing recordings, particularly of *melodies*. This allows the theory to propose that the most common scale-degree sizes—those usually smaller than a fourth—may be schematically representative of the small steps used in melodic motion (Dowling 1978). This creates an intrinsic connection between scales and melodies. There are additional reasons for a scale theory to prioritize melody over harmony. The majority of global music traditions are either primarily melodic or primarily rhythmic (Rich 2000). Accounting for the relatively understudied properties of melodic scales (c.f. Panteli, Benetos, and Dixon 2018) should be an important objective in this field.

Of course, harmony and polyphony *have* exerted a significant influence on the development of certain scales, especially those of the Western art canon. Pythagorean "perfect" intervals are ideal for creating harmonic textures that minimize the sensory percept of overtone beating, especially in large resonant spaces such as cathedrals, where large pipe organs capable of producing such intervals with high replicability could be engineered and installed. The "consonant" effect of minimal beating, combined with a particularly pure vocal timbre, contributed to the crystallization of a European polyphonic choral and organ repertoire—known as organum—that was full of "perfect" intervals, at least at the principal cadential points. Many European composers throughout the millennia have designed their harmonic motion around perfect intervals to indicate that they were mathematically literate and sensitive to the "ideal construction principles" dominating most of Western art music (Maróthy and Norton 1985).

Scales influenced by specific harmonic considerations are by no means unique to the Western canon. Many of the world's art musics have rich harmonic textures, especially those based on strings and pipes. For example, Persian classical music was primarily developed by the social elite of the seventh-century Sasanian empire, notably the court bard Barbad. Barbad is credited with the invention of the lute as well as the seven "royal modes," 30 derivative modes, and 360 of the core melodies of Iran's classical music. These were developed in the following centuries by Muslim mathematicians and philosophers, including Avicenna (Matrasul 2017; Parvin and Afkhami

2020). This flourishing art music was driven, just as in Western Europe, by the high-society's intellectual fascination with the links between harmony, mathematics, and instrumental engineering. Of course, harmony is also central to many traditional and popular musics, including indigenous choral traditions like the well-known *haka* polyphony of the Central African Aka pygmies (Arom 1991, Arom, and Fürniss 1993; Fürniss 2006; Wheeler 2018), although many cultures eschew formalized rules and labels for these harmonies (Ekwueme 1974; Toub 1999).

It must be noted, however, that harmony is not a universal element across world music traditions, whereas some form of melody is necessarily present in all tonal music, even those that emphasize harmony as well. In evolutionary terms, a solo voice must necessarily predate multiple voices, at least in songs with organized tonality. Interval Spacing theory thus investigates melodic, rather than harmonic, scales on the basis of their cultural primacy and generalizability. The theory is likewise based on scale cognition, rather than mathematical theory, because (melodic) scales are more fundamental constructs than modes, keys, and tonal hierarchies.

Recall that a scale is simply a set of pitch classes, or more accurately a set of intervals or pitch-class spacings. Scales need not be anchored to any specific pitch, nor related to any particular melody. They do not, on their own, imply harmonic relationships, hierarchies, or voice-leading rules. The only function of a scale is to describe a common set of melodic notes, or to provide a scaffold of common notes for generating future melodies. This definition may be counterintuitive to the Western music scholar, precisely because our most common scales are actually medieval modes. Modes, in comparison to scales, regulate the *order* of intervals by labeling one primary note as the "tonic" or "final."

Take the major scale (or Dorian mode) as an example. It is commonly taught as a pattern of whole steps (w) and half steps (h), that is, w-w-h-w-w-w-h, progressing from a given note. If a modal melody spans more than an octave, outer notes can be transposed up or down as necessary in order to fit the melody within the octave-bound interval pattern. Modes are thus more complex constructs than scales. Whereas scales simply describe the salient notes of a melody, modes compare the notes across octaves, collapse them into categorical chroma, and rearrange them according to the primacy of one particular pitch class, a process that requires information from prior experience with melodies composed in that mode.

Keys and tonal hierarchies require an even greater extent of prior musical knowledge, because they also rank each note within a hierarchy of importance relative to the tonic. These hierarchies, unlike the more immediate and universal cognitive representations of scales (Dowling 1978), derive from long-term listening experience within a certain musical culture (Hébert, Peretz, and Gagnon 1995; Burns 1999) and therefore provide a wealth of contextual information about the melody. They allow listeners to predict which notes will co-vary with other notes, which notes will likely follow other notes, which keys are closely related to one another and may be pivoted to, and even the possible functions of chromatic notes that would otherwise be deviant tones within a scale (Justus and Bharucha 2002).

Such context-based listening, starting with the perception of octave equivalence and a tonic note, seems to activate automatically among listeners with the relevant musical experience. This fact can make it difficult for Western music scholars to "unhear" keys and modes and to instead conceptualize scales in the most basic melodic sense described above. However, from an evolutionary and cross-cultural point of view, sequential intervals are both more ancient and universal than simultaneous intervals, melody is more ancient and universal than harmony, and complex structures tend to develop over time from simpler progenitors. Studies about key perception, the diatonic key system, and 12-ET (Krumhansl and Kessler 1982; Krumhansl 1991; Krumhansl 2004) reveal more about texture and harmony in particular cultures than they do about *scales* in a cognitively general, evolutionarily relevant sense.

An account of such scales, grounded in heuristic descriptions of melodies, has not been well-developed thus far. Such an account could provide stable fundamentals for building more-complex theories of more-complex tonal schemata. It could also motivate research into the cognitive "transfer functions" that convert melody representations into scale schemata, or that replace scales with modes and keys (Matsunaga and Abe 2005, 2007). This view of scales would also emphasize the fundamental roles of enculturation and musical experience in these cognitive processes, in contrast to scale theories that presume that the patterns seen in Western music (and their consequent cognitive processes) are the most natural, most common, or most mathematically perfect of patterns.

Why a vocal theory?

Interval Spacing theory's prioritization of what is primary, ancient, and universal also leads to an emphasis on vocal scales above instrumental scales. We have mentioned that art musics, both Western and otherwise, are usually influenced by mathematical theories of harmonic ratios. Such theories are almost always based on tunable instruments, which have starkly different intonational precision and accuracy compared to the unaccompanied voice, and even to unaccompanied choruses. Vocal music has no *a priori* tuning, consists of highly imprecise notes and intervals relative to fixed-pitch instruments (Pfordresher et al. 2010; Pfordresher and Brown 2017), and is full of complex melodic and intonational gestures. Therefore, vocal scales ought to be highly complex and variable, and a scale theory that can account for these complexities should be well-equipped to subsequently model instrumental scales with preset tunings and relatively straightforward pitch patterns.

The voice is the oldest and most universal instrument. The most compelling theories about the evolution of scales should emanate from an analysis of the structural features of vocal scales and of relating these features to potential sensorimotor and memory constraints. This may be especially true in cultures with a longstanding tradition of oral musical transmission, where the influence of written theory and instrumental practice on vocal music is more limited or even non-existent. The commonalities among globally diverse vocal traditions ought to be related to relatively universal physiological and cognitive explanations, although culture-specific factors will influence the melodic content of any musical tradition. The more purely vocal a musical tradition is, the more valid these physiological explanations will be. Interval Spacing theory encourages the analysis of vocal melodies for three reasons: first, to ensure that the raw data regarding scale structure are as unbiased as possible; second, to ensure that the methods of descriptive scale analysis are sufficiently powerful to tackle even the most complex pitch patterns; and third, to make compelling arguments about the origin of musical scales.

Towards a Combinatorics of Scale Generation

Having provided the justification for a theory of musical scales that is vocal and melodic, we conclude this chapter by attempting to resolve the chicken-and-egg problem that arises from a descriptive, rather than prescriptive, model of musical scales. How can scales emerge without musicians having theoretical *a priori* notions of the pitch classes to begin with, such as those derived from instrumental tunings (see Figure 2)? How exactly can scales be related to melodic motion? We believe that the mechanisms of generativity, employing combinations of a few basic interval building-blocks, may be the key. Sachs (1943) proposed an evolutionary model in which scales were built up over historical time, from the earliest monotonic chants to two-note scales, three-note scales, and beyond. To model this development, one has to constrain the potential distance at which new pitch classes can be added to a previous pitch class (or pitch class set) to expand the scale. Sachs refers to such additions as "affixes." For example, a two-note scale can become a three-note scale by affixing a pitch to the upper pitch of the initial scale (or alternatively to its lower pitch).

We previously discussed the tenet of the Interval Spacing theory that the whole tone (200 cents), rather than the semitone, is the more vocally distinguishable interval, and hence the more reliable interval for scale generation. However, the favorability of the whole tone scale with respect to vocal spacing is offset by the fact that the majority of scales show a strong preference for non-equidistance. Having more than one interval size in a scale creates asymmetry in the scale and thereby permits the existence of a tonic pitch and an associated tonal hierarchy, which are usually absent from equidistant scales (although see Ross and Knight 2019). Therefore, instead of the whole tone scale, we see a generative principle by which scales are built up of sequential combinations of intervals that are "a whole tone +/- a semitone." In Western terms, this means semitones, whole tones, and minor thirds, although we are not implying that scales can *only* consist of steps that are 100, 200, and 300 cents exactly, as may be posited by Harmonicity theory. Rather, scale step sizes generally fall anywhere within the *range* of 100-300 cents. While the lower limit is constrained by vocal imprecision, the upper limit is more flexible, where step sizes as large as 500 cents are known to occur in many global scales, and even larger steps are used occasionally. In referring to these constraints on scale step sizes, we are not discounting the presence or importance of melodic gestures and ornaments smaller than a semitone.

Figure 4 shows that the Interval Spacing model provides novel insight into salient properties of musical scales, including the minimum step size (a semitone), the inventory of step sizes (most scales have only 2-3 step types), and the number of scale tones (typically 5-7/octave, although a scale need not contain 5 or more tones nor span the range of an octave). It is critical to consider that none of these features are predictable from Harmonicity theory and its *a priori* restrictions. Rather, by placing the evolution of scales firmly within the realm of combinatorics, the Interval Spacing theory generates a wealth of such observable—and testable—predictions.

In parallel to the standard notion of musical combinatorics, by which the pitch classes of a scale are used as building blocks to form melodies, we propose that *scales themselves are built up as sequential combinations of various step sizes*, such as semitones, whole tones and minor thirds. From the combinations of a small number of such building blocks, the huge diversity of world scales is possible. Sachs described this well when he said that scale combinatorics results in “a kaleidoscopic infinity of variations and permutations” (1943, 39). Building scales combinatorially from the bottom up provides a promising alternative to Harmonicity theory’s top-down approach, where the justly intoned pitch scaffold has no flexibility and allows for no deviations.

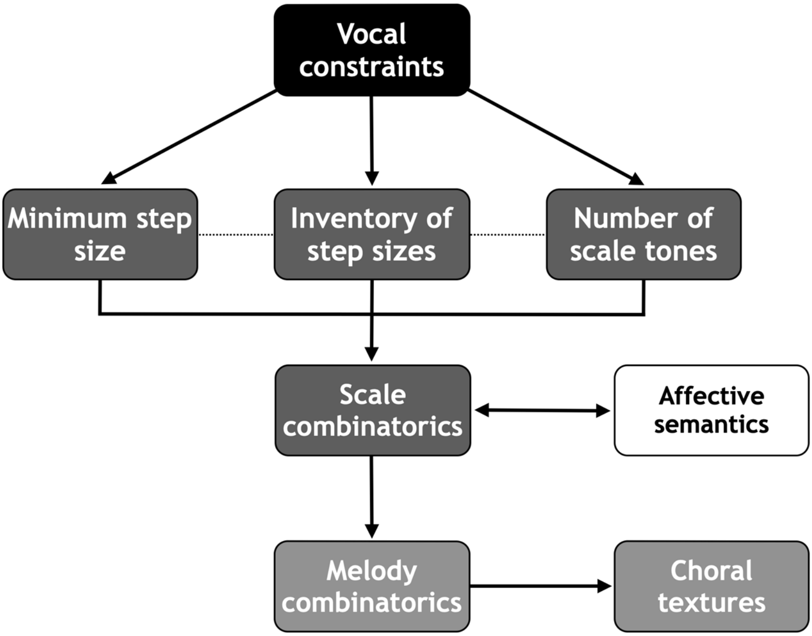

Figure 4: A model for the combinatoric generation of scales and melodies based on the Interval Spacing theory, combined with a consideration of “affective semantics” related to the communicative meaning of different scale types. Choral textures, such as chordal harmony, are derived from a combination of melodies.

A global analysis of the interval combinations that are favored (or disfavored) in the structure of world scales, most especially of vocal scales, should provide a means of deriving some generative "rules" and patterns for scale combinatorics. For example, vocal constraints suggest that two adjacent semitones should be disfavored in musical scales. Indeed, outside of the Western chromatic scale, it is extremely rare to find scales having adjacent semitones. We are currently engaged in a global analysis of vocal scales in traditional cultures in order to discover other such regularities in the structure of vocal scales. While it is too early to report on the results, we believe that these findings will form the basis for a new theory of the origin of musical scales (Figure 4). That theory will operate based on bottom-up principles of scale combinatorics, grounded in vocal-motor factors that both constrain the sizes of adjacent intervals and that favor certain combinations of interval sequences over others. A critical factor that will affect the adoptability of a given scale in a particular culture will be the affective meaning of the scale for its users. Hence, the physiology of interval spacings will ultimately need to interface with cognitive factors that determine the "affective semantics" of a musical scale and how that scale communicates meaningful information to members of a culture (Brown 2017). Transmissibility, and its related memory constraints, will also play a role in the cultural evolution of musical scales.

Finally, we bring the model full circle back to harmony, which we have left unaccounted for thus far. We posit that musical harmony and texture also have the features of a combinatorial system. They are based on how individual melodic lines are joined together in both time and tonal pitch-space to create multipart combinations (Huron 1989; Jordania 2006). Therefore, harmony could have originally arisen from the joining together of vocal lines to create choral textures through interpersonal entrainment mechanisms, not from the harmonics of single pitched sounds (Brown 2007). As such, harmony is the *last* step in our model of musical scale evolution, whereas it is the very groundwork of Harmonicity theory. Figure 4 summarizes the three distinct mechanisms by which combinational mechanisms generate pitch structure in music. Musical scales are proposed to be derived from a mechanism of "affixation" by which basic building blocks, constrained by physiological mechanisms, are combined together into a linear scale sequence. Melodies can then be generated by a combination of scale pitches. Finally, harmonies and textures can be generated by a combination of melodies.

Conclusions

Although Harmonicity theory has been the historically dominant lens through which scholars have studied scale structure, we argue that shifting the focus away from *a priori* theory and physics towards physiology and culture offers a better descriptive and explanatory model of musical scales. We developed one such theory called Interval Spacing theory. It is derived from the observation that vocal pitch and interval classes are imprecise and categorical, suggesting a division of pitch space into "stepping stones" or "island" patterns, rather than exact ladder rungs. The model's fundamental proposition is that the structural patterns of many scales represent a trade-off between interval distinguishability and interval producibility. Scale pitches must be far enough apart that their distributions are largely separate, but not so far that the interval is overly difficult or undesirable to sing. This single rule, based on the pitch-production constraints of the voice, suggests a few basic building blocks for melodic pitches and intervals, which can be combined into complex patterns to produce the wealth of vocal melodies and scales heard around the world.

To evaluate this theory—or to evaluate any theory of scales—we argue for the analysis of a culturally diverse sample of music, developing empirical (but musicologically relevant) descriptive models thereof, and investigating how those models may be linked to physiological, cognitive, cultural, and evolutionary evidence. Such an account of music, which is focused on vocal performance, is not without precedent. For millennia, scholars have studied song for its intrinsic connection with speech and emotion, its primacy, and its universality. This focus, however, has not yet penetrated the theory- and philosophy-dominated scholarship surrounding musical scales in a significant way. We hope that Interval Spacing theory will further promote the unbiased study of vocal melodies, especially those maintained and transmitted in longstanding oral traditions, and thereby expand the scope of the field. In tandem, we hope that the theory will offer some new explanatory hypotheses for pitch and scale structure based on empirical evidence. At the very least, we hope that it will challenge the notion that the hypotheses of Harmonicity theory are irrefutable facts of nature and are the sole governance of patterns in human music.

Acknowledgments

This work was supported by a grant to S.B. from the Social Sciences and Humanities Research Council (SSHRC) of Canada (grant number 435-2017-0491).

Bibliography

Ahlbäck, Sven. 2004. "Melody Beyond Notes: A Study of Melody Cognition." Doctoral diss., Skrifter Från Institutionen För Musikvetenskap, Göteborgs Universitet, nr 77. Göteborgs universitet.

Arom, Simha. 2004. *African Polyphony and Polyrhythm: Musical Structure and Methodology*. Cambridge: Cambridge University Press.

Arom, Simha, and Susanne Fürniss. 1993. "An Interactive Experimental Method for the Determination of Musical Scales in Oral Cultures: Application to the Vocal Music of the Aka Pygmies of Central Africa." *Contemporary Music Review*. Vol. 9, No. 1-2. 7-12.

Batteux, Charles. 1745/2015. *The Fine Arts Reduced to a Single Principle.* Translated by J. O. Young. Oxford: Oxford University Press.

Berger, Lawrence M. 2000. "The Emotional and Intellectual Aspects of Protest Music: Implications for Community Organizing Education." *Journal of Teaching in Social Work*. Vol. 20, No. 1-2. 57-76.

Bharucha, Jamshed J., and Peter M. Todd. 1989. "Modeling the Perception of Tonal Structure with Neural Nets." *Computer Music Journal*. Vol. 13, No. 4. 44-53.

Brown, Steven. 2007. "Contagious Heterophony: A New Theory About the Origins of Music." *Musicae Scientiae*. Vol. 11, No. 1. 3-26.

Brown, Steven. 2017. "A Joint Prosodic Origin of Language and Music." *Frontiers in Psychology*. Vol. 8. 1894.

Burns, Edward M., and W. Dixon Ward. 1978. "Categorical Perception—Phenomenon or Epiphenomenon: Evidence from Experiments in the Perception of Melodic Musical Intervals." *The Journal of the Acoustical Society of America*. Vol. 63, No. 2. 456-468.

Burns, Edward M. 1999. "Intervals, Scales, and Tuning." In *The Psychology of Music,* edited by Diana Deutsch, 2nd ed., 215-264. San Diego: Academic Press.

Cartwright, Julyan H. E., Diego L. González, Oreste Piro, and Domenico Stanzial. 2002. "Aesthetics, Dynamics, and Musical Scales: A Golden Connection." *Journal of New Music Research*. Vol. 31, No. 1. 51-58.

Cazden, Norman. 1945. "Musical Consonance and Dissonance: A Cultural Criterion." *The Journal of Aesthetics and Art Criticism*. Vol. 4, No. 1. 3-11.

Cazden, Norman. 1980. "The Definition of Consonance and Dissonance." *International Review of the Aesthetics and Sociology of Music*. Vol. 11, No. 2. 123-168.

Chiba, Gakuto, Meng-Jou Ho, Shoichiro Sato, Jiei Kuroyanagi, Joren Six, Peter Pfordresher, Adam Tierney, Shinya Fujii, and Patrick E. Savage. 2019. "Small-integer Ratios Predominate Throughout the World's Musical Scales." Preprint, *PsyArXiv*.

Cross, Ian, Peter Howell, and Robert West. 1983. "Preferences for Scale Structure in Melodic Sequences." *Journal of Experimental Psychology: Human Perception and Performance*. Vol. 9, No. 3. 444-460.

de Cheveigné, Alain, and Hideki Kawahara. 2002. "YIN, a Fundamental Frequency Estimator for Speech and Music." *The Journal of the Acoustical Society of America*. Vol. 111, No. 4. 1917-1930.

Deutsch, Diana. 1969. "Music Recognition." *Psychological Review*. Vol. 76, No. 3. 300-307.

Dowling, W. Jay. 1978. "Scale and Contour: Two Components of a Theory of Memory for Melodies." *Psychological Review*. Vol. 85, No. 4. 341-354.

Dowling, W. Jay. 2002. "The Development of Music Perception and Cognition." In *Foundations of Cognitive Psychology: Core Readings,* edited by Daniel J Levitin, 481-502. Cambridge: MIT press.

Dowling, W. Jay, and James C. Bartlett. 1981. "The Importance of Interval Information in Long-term Memory for Melodies." *Psychomusicology*. Vol. 1, No. 1. 30-49.

Durfee, Dallin S., and John Snyder Colton. 2015. "The Physics of Musical Scales: Theory and Experiment". *American Journal of Physics*. Vol. 83, No. 10. 835-842.

Ekwueme, Laz E. N. 1974. "Concepts of African Musical Theory." *Journal of Black Studies*. Vol. 5, No. 1. 35-64.

Ellis, Alexander John. 1885. "On the Musical Scales of Various Nations." *Journal of the Society of Arts*. Vol. 33. 485-527.

Ellis, Catherine J. 1965. "Pre-instrumental Scales." *Ethnomusicology*. Vol. 9, No. 2. 126-137.

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

Fürniss, Susanne. 2006. "Aka Polyphony: Music, Theory, Back and Forth." In *Analytical Studies in World Music,* edited by M. Tenzer, 163-204. Oxford: Oxford University Press.

Geringer, John M., Rebecca B. MacLeod, Clifford K. Madsen, and Jessica Napoles. 2015. "Perception of Melodic Intonation in Performances With and Without Vibrato." *Psychology of Music*. Vol. 43, No. 5. 675-685.

Geringer, John M., and David W. Sogin. 1988. "An Analysis of Musicians' Intonational Adjustments Within the Duration of Selected Tones." *Contributions to Music Education*. Vol. 15. 1-6.

Gill, Kamraan Z., and Dale Purves. 2009. "A Biological Rationale for Musical Scales." *PLOS ONE*. Vol. 4, No. 12. e8144.

Hall, Donald E., and Joan Taylor Hess. 1984. "Perception of Musical Interval Tuning." *Music Perception*. Vol. 2, No. 2. 166-195.

Hébert, Sylvie, Isabelle Peretz, and Lise Gagnon. 1995. "Perceiving the Tonal Ending of Tune Excerpts: The Roles of Pre-existing Representation and Musical Expertise." *Canadian Journal of Experimental Psychology*. Vol. 49, No. 2. 193-210.

Helmholtz, Hermann L. F. 1885. "Mathematical and Physical Papers." *Nature*. Vol. 32, No. 811. 25-28.

Honingh, Aline, and Rens Bod. 2011. "In Search of Universal Properties of Musical Scales." *Journal of New Music Research*. Vol. 40, No. 1. 81-89.

Huron, David. 1989. "Voice Denumerability in Polyphonic Music of Homogeneous Timbres." *Music Perception*. Vol. 6, No. 4. 361-382.

Huron, David. 1994. "Interval-class Content in Equally Tempered Pitch-class Sets: Common Scales Exhibit Optimum Tonal Consonance." *Music Perception*. Vol. 11, No. 3. 289-305.

Huron, David. 2004. "Issues and Prospects in Studying Cognitive Cultural Diversity." In *Proceedings of the 8th International Conference on Music Perception and Cognition*, edited by S. Lipscomb, R. Ashley, R. Gjerdingen, and P. Webster, 93-95. Evanston, IL: The Society for Music Perception and Cognition.

Hutchins, Sean, Catherine Roquet, and Isabelle Peretz. 2012. "The Vocal Generosity Effect: How Bad Can Your Singing Be?" *Music Perception*. Vol. 30, No. 2. 147-159.

Hutchinson, William, and Leon Knopoff. 1978. "The Acoustic Component of Western Consonance." *Journal of New Music Research*. Vol. 7, No. 1. 1-29.

Jordan, Daniel S., and Roger N. Shepard. 1987. "Tonal Schemas: Evidence Obtained by Probing Distorted Musical Scales." *Perception & Psychophysics*. Vol. 41, No. 6. 489-504.

Jordania, Joseph. 2006. *Who Asked the First Question? The Origins of Human Choral Singing, Intelligence, Language and Speech.* Tblisi, Georgia: Logos.

Justus, Timothy C., and Jamshed J. Bharucha. 2002. "Music Perception and Cognition." In *Stevens' Handbook of Experimental Psychology:*

Sensation and perception, edited by H. Pashler and S. Yantis, 453-492. Hoboken: John Wiley & Sons.

Kameoka, Akio, and Mamoru Kuriyagawa. 1969. "Consonance Theory Part I: Consonance of Dyads." *The Journal of the Acoustical Society of America*. Vol. 45, No. 6. 1451-1459.

Kenstowicz, Michael, and Charles Kisseberth. 2014. *Generative Phonology: Description and Theory*. New York: Academic Press.

Killick, Andrew. 2020. "Global Notation as a Tool for Cross-cultural and Comparative Music Analysis." *Analytical Approaches to World Music*. Vol. 8, No. 2. 235-279.

Kristeller, Paul Oskar. 1951. "The Modern System of the Arts: A Study in the History of Aesthetics Part I." *Journal of the History of Ideas*. Vol. 12, No. 4. 496-527.

Krumhansl, Carol L. 1991. "Music Psychology: Tonal Structures in Perception and Memory." *Annual Review of Psychology*. Vol. 42, No. 1. 277-303.

Krumhansl, Carol L., and Edward J. Kessler. 1982. "Tracing the Dynamic Changes in Perceived Tonal Organization in a Spatial Representation of Musical Keys." *Psychological Review*. Vol. 89, No. 4. 334-368.

Kuroyanagi, Jiei, Shoichiro Sato, Meng-Jou Ho, Gakuto Chiba, Joren Six, Peter Pfordresher, Adam Tierney, Shinya Fujii, and Patrick E. Savage. 2019. "Automatic Comparison of Human Music, Speech, and Bird Song Suggests Uniqueness of Human Scales." *Proceedings of the 9th International Workshop on Folk Music Analysis*. 35–40.

Li, Yipeng, and David Huron. 2006. "Melodic Modeling: A Comparison of Scale Degree and Interval." *ICMC.* 495-498.

Loosen, Franz. 1994. "Tuning of Diatonic Scales by Violinists, Pianists, and Nonmusicians." *Perception & Psychophysics*. Vol. 56, No. 2. 221-226.

Loosen, Franz. 1995. "The Effect of Musical Experience on the Conception of Accurate Tuning." *Music Perception*. Vol. 12, No. 3. 291-306.

Maróthy, János, and Richard Norton. 1985. "Tonality in Western Culture: A Critical and Historical Perspective." *Studia Musicologica Academiae Scientiarum Hungaricae*. Vol. 27, No. 1/4. 400.

Martin, Daniel W. 1962. "Musical Scales Since Pythagoras." *Sound: Its Uses and Control*. Vol. 1, No. 3. 22-24.

Mason, James A. 1960. "Comparison of Solo and Ensemble Performances with Reference to Pythagorean, Just, and Equi-tempered Intonations." *Journal of Research in Music Education*. Vol. 8, No. 1. 31-38.

Matrasul, Matyakubov. 2017. "From History Music Performance Khorezm." *European Journal of Arts*. Vol. 1. 21-24.

Matsunaga, Rie, and Jun-Ichi Abe. 2005. "Cues for Key Perception of a Melody: Pitch Set Alone?" *Music Perception*. Vol. 23, No. 2. 153-164.

Matsunaga, Rie, and Jun-Ichi Abe. 2007. "Incremental Process of Musical Key Identification." *Proceedings of the 29th Annual Meeting of the Cognitive Science Society*. 1277-1282.

Mauch, Matthias, Klaus Frieler, and Simon Dixon. 2014. "Intonation in Unaccompanied Singing: Accuracy, Drift, and a Model of Reference Pitch Memory." *The Journal of the Acoustical Society of America*. Vol. 136, No. 1. 401-411.

Mcbride, John, and Tsvi Tlusty. 2019. "Imperfect Fifths Pack into Musical Scales." Preprint, *PsyArXiv*.

McDermott, Josh H., Alan F. Schultz., Eduardo A. Undurraga, and Ricardo A. Godoy. 2016. "Indifference to Dissonance in Native Amazonians Reveals Cultural Variation in Music Perception." *Nature*. Vol. 535, No. 7613. 547-550.

Morrison, Steven J. 2000. "Effect of Melodic Context, Tuning Behaviors, and Experience on the Intonation Accuracy of Wind Players." *Journal of Research in Music Education*. Vol. 48, No. 1. 39-51.

Ortmann, Otto. 1926. "On the Melodic Relativity of Tones." *Psychological Monographs*. Vol. 35, No. 1. 1–47.

Panteli, Maria, Emmanouil Benetos, and Simon Dixon. 2017. "A Computational Study on Outliers in World Music." *PLOS ONE*. Vol. 12. e0189399.

Panteli, Maria, Emmanouil Benetos, and Simon Dixon. 2018. "A Review of Manual and Computational Approaches for the Study of World Music Corpora." *Journal of New Music Research*. No. 47, No. 2. 176-189.

Pantev, Christo, Thomas Elbert, Bernhard Ross, Carsten Eulitz, and Ernst Terhardt. 1996. "Binaural Fusion and the Representation of Virtual Pitch in the Human Auditory Cortex." *Hearing Research*. Vol. 100, No. 1-2. 164-170.

Parncutt, Richard, and Graham Hair. 2018. "A Psychocultural Theory of Musical Interval: Bye Bye Pythagoras." *Music Perception: An Interdisciplinary Journal*. Vol. 35, No. 4. 475-501.

Parvin, Samad, and Behrouz Afkhami. 2020. "Sasanid Music (From Historical Texts to Archaeological Evidence)." *Вестник Томского Государственного Университета. Культурология и Искусствоведение*. Vol. 37. 165-174.

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

Pfordresher, Peter Q., and Steven Brown. 2017. "Vocal Mistuning Reveals the Origin of Musical Scales." *Journal of Cognitive Psychology*. Vol. 29, No. 1. 35-52.

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*. Vol. 128, No. 4. 2182-2190.

Rakowski, Andrzej. 1990. "Intonation Variants of Musical Intervals in Isolation and in Musical Contexts." *Psychology of Music*. Vol. 18, No. 1. 60-72.

Rameau, Jean-Philippe. 1722/1971. *Treatise on Harmony*. Translated by Philip Gossett. New York: Dover.

Rich, Alan. 2000. "Harmony." In *Encyclopedia Britannica*. https://www.britannica.com/art/harmony-music

Ross, Barry, and Sarah Knight. 2019. "Reports of Equitonic Scale Systems in African Musical Traditions and Their Implications for Cognitive Models of Pitch Organization." *Musicae Scientiae*. Vol. 23, No. 4. 387–402.

Rousseau, Jean-Jacques. 1781/1998. *Essay on the Origin of Languages*. Translated and edited by John T. Scott. Hannover, NH: University Press of New England.

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

Savage, Patrick E., Steven Brown, Emi Sakai, and Thomas E. Currie. 2015. "Statistical Universals Reveal the Structures and Functions of Human Music." *Proceedings of the National Academy of Sciences*. Vol. 112, No. 29. 8987-8992.

Schellenberg, E. Glenn, and Laurel J. Trainor. 1996. "Sensory Consonance and the Perceptual Similarity of Complex-tone Harmonic Intervals: Tests of Adult and Infant Listeners." *The Journal of the Acoustical Society of America*. Vol. 100, No. 5. 3321-3328.

Schellenberg, E. Glenn, and Sandra E. Trehub. 1996. "Natural Musical Intervals: Evidence from Infant Listeners." *Psychological Science*. Vol. 7, No. 5. 272–277.

Schmuckler, Mark A. 2009. "Components of Melodic Processing." In *Oxford Handbook of Music Psychology*, edited by Susan Hallam, Ian Cross, and Michael Thaut, 93-106. Oxford: Oxford University Press.

Siegel, Jane A., and William Siegel. 1977a. "Absolute Identification of Notes and Intervals by Musicians." *Perception & Psychophysics*. Vol. 21, No. 2. 143-152.

Siegel, Jane A., and William Siegel. 1977b. “Categorical Perception of Tonal Intervals: Musicians Can’t Tell Sharp from Flat.” *Perception & Psychophysics*. Vol. 21, No. 5. 399-407.

Siegel, William, and Robert Sopo. 1975. “Tonal Intervals are Perceived Categorically by Musicians with Relative Pitch.” *The Journal of the Acoustical Society of America*. Vol. 57. S11.

Terhardt, Ernst. 1984. “The Concept of Musical Consonance: A Link Between Music and Psychoacoustics.” *Music Perception*. Vol. 1, No. 3. 276-295.

Terhardt, Ernst. 1991. “Music Perception and Sensory Information Acquisition: Relationships and Low-level Analogies.” *Music Perception*. Vol. 8, No. 3. 217-239.

Thomas, Downing A. 1995. *Music and the Origins of Language: Theories from the French Enlightenment*. Cambridge: Cambridge University Press.

Toub, Martin D. 1999. *Harmony* (Vol. 9). Pacific Street Films and the Educational Film Center. https://www.learner.org/series/exploring-the-world-of-music/harmony/

Vos, Piet G., and Jim M. Troost. 1989. “Ascending and Descending Melodic Intervals: Statistical Findings and Their Perceptual Relevance.” *Music Perception*. Vol. 6, No. 4. 383-396.

Vurma, Allan, and Jaan Ross. 2006. “Production and Perception of Musical Intervals.” *Music Perception*. Vol. 23, No. 4. 331-344.

Ward, W. Dixon. 1970. “Musical Perception.” In *Foundations of Modern Auditory Theory*, edited by Jerry V. Tobias, 407-447. New York: Academic Press.

Warren, Richard A., and Maegan E Curtis. 2015. “The Actual vs. Predicted Effects of Intonation Accuracy on Vocal Performance Quality.” *Music Perception*. Vol. 33, No. 2. 135-146.

Wheeler, Paul A. 2018. “Traditional Polyphony: Multipart Singing in World Cultures.” *The Journal of the Acoustical Society of America*. Vol. 143, No. 3. 1842.

Zatorre, Robert J., and Andrea R. Halpern. 1979. “Identification, Discrimination, and Selective Adaptation of Simultaneous Musical Intervals.” *Perception & Psychophysics*. Vol. 26, No. 5. 384-395.