# Hierarchizability as a Predictor of Scale Candidacy

Niels J. Verosky
*San Francisco, CA*

**Scales from divergent musical cultures tend** to have both intuitive structural similarities and one common functional property: within a given scale, each note takes on a unique shade of meaning in the context of the scale as a whole. It may be that certain structural traits facilitate this functional property—in other words, that scales with particular structural characteristics are more globally integrated and capable of being processed in a top-down manner. Representing pitch collections as bit strings, the current work shows that in Western European, Northern Indian, and Japanese traditional musics, collections that are more densely packed with recursively nested non-overlapping, uniquely identifiable repeated substrings (more *hierarchizable*) are more likely to appear as scales ($p = .002$).

*Received: June 23, 2016, accepted November 7, 2016.*

**Key words:** scales, scale candidacy, properties of scales, universal constraints on music, hierarchy

There are enough intuitive similarities between scales from divergent musical cultures to suggest strong universal constraints on scale candidacy, but these similarities are on the surface fluid enough as to resist precise description. For instance, many common scales in world music combine larger and smaller steps in a way that creates obvious parallelism but introduces a slight asymmetrical twist (e.g., the pentatonic's 3 + 2 grouping). Similarly, scales taken from musical traditions around the world tend to favor more or less even distributions of notes across the octave, yet transpositionally symmetrical (T-symmetrical) scales like the whole-tone collection—the most even of all—are rare. Seven- and five-note scales are vastly overrepresented in comparison with six-note scales, and while one approach to this observation points out that seven- and five-note collections in 12-pc space tend to have the "symmetrical but not too symmetrical" structures typical of scales in world music, this explanation merely recasts the question of why five- and seven-note collections predominate in terms of the question of why slightly asymmetrical scales seem to be preferred. In general, these mixed results—where positing an answer to one open question highlights the lack of satisfying answers to other open questions—are typical of theories isolating certain similarities between scales taken from different cultures.

Beyond their common structural attributes, scales across world music largely share one functional attribute: collections used as scales are not merely flat fields from which pitches are drawn independently, but rather tend to act as globally integrated totalities, such that each individual pitch takes on a unique shade of meaning within the context of the scale as a whole. In the familiar major scale, for instance, probe tone studies by Krumhansl and Kessler (1982) suggest that Western listeners hear the tonic as more congruent with the overall key than the fifth, the fifth as more congruent than the third, the third as more congruent than the fourth, and so on. As Krumhansl and Cuddy (2010) point out, other studies have found similar results for Indian and Balinese listeners, indicating that scales serve as webs of pitch relations (as opposed to mere collections of pitches) across cultures even if the particular relations they are assigned in musical practice are culture-specific (Castellano, Bharucha, & Krumhansl, 1984; Kessler, Hansen, & Shepard, 1984).

Given this culture-to-culture functional similarity between scales, the question arises as to whether some pitch collections are more easily able to assume these kinds of unified orderings than others. If pitch structures with certain properties appear more frequently as scales across world music, it may be that these pitch structures are better suited to storing information about individual pitches' global significance in terms of the scale as a whole than pitch structures that appear rarely or not at all. The current study will suggest that more *hierarchizable* pitch collections—that is, pitch collections exhibiting certain traits that allow them to more easily be parsed in top-down, differentiated, cohesive terms—are disproportionately represented throughout world music.

Because *hierarchizability* (the extent to which a given pitch set is capable of being assigned this kind of globally integrated, hierarchical structure) as presented below represents a well-defined structural property with a strong corresponding functional interpretation, the

*Music Perception*, volume 34, issue 5, pp. 515–530, issn 0730-7829, electronic issn 1533-8312.  DOI: https://doi.org/10.1525/mp.2017.34.5.515

idea of hierarchizability underscores a potential bridge between structural constraints (the "regular but not too regular" quality exemplified by scales like the diatonic, harmonic minor, and *pelog*) and functional constraints (the fact that individual pitches within scales play unique roles within the context of the entire scale) apparent in scales across world music. The structural side of this bridge especially involves previous work exploring constraints on scale candidacy—in particular, work on evenness and pattern matching as determinants of scale candidacy.

Perhaps the most empirically convincing demonstration of evenness as a universal criterion for scale candidacy appears in Carey (2002). Carey ranks *n*-pc set-classes by "coherence quotient"—in general terms, a measure of a specific kind of evenness—and shows that across several cultures, common scales tend to be drawn from the most coherent 5- and 7-note sets. However, because T-symmetrical scales are the most coherent of all, Carey's coherence quotient most strongly predicts scale candidacy under the a priori assumption that 5- and 7- but not 6-note scales are the best scale candidates, indicating that either (a) coherence quotient interacts with at least one other significant variable to predict scale candidacy, or (b) coherence quotient and possibly evenness in general are not themselves primary determinants of scale candidacy but rather are affected by some other property that more strongly predicts scale candidacy.

Browne's work on pattern matching and position finding, while less empirically oriented than Carey's, suggests intuitively compelling properties that might contribute to scale candidacy (Browne, 1981). In particular, his emphasis on the interdependence of matchable patterns and position finding intervals that clarify global context fits with the "symmetrical but not too symmetrical" nature of many common scales. Nonetheless, while Browne provides a strong argument for the widespread popularity of the diatonic, it is less clear how his concepts relate to other scales that recur culture to culture across world music.

The following will suggest that recasting pattern matching, position finding, content, and context in terms of scale fragments rather than intervals gives Browne's ideas a more universal relevance, and that insofar as pattern matching along the lines of scale fragments implies a degree of evenness, evenness as a criterion for scale candidacy becomes subsumed under pattern matching and position finding as determinants of scale candidacy. Experimental work has shown that people are sensitive to repeating patterns in tone sequences (Deutsch, 1980; Herholz, Lappe, & Pantev, 2009), melodies (Volk & van Kranenburg, 2012), and even randomly generated visual inputs (Huettel, Mack, & McCarthy, 2002), raising the possibility that listeners may use matchable patterns to organize information about pitch collections and infer scales as well.

Like pattern matching, position finding connects with previous work on music perception. Trehub, Schellenberg, and Kamenetsky (1999) found that infants were better at detecting mistuned notes in the major scale and in a novel scale with unequal steps than in a novel scale with equal steps, while adults were best at detecting mistuned notes in the major scale and equally bad at detecting deviations in novel scales with both equal and unequal steps. This result indicates that position finding may be especially salient in scale acquisition.

To simplify stepwise pattern-matching operations on scale fragments, the study that follows tends towards representing scales not in terms of pitch-class set-classes but in terms of equivalent bit string classes. The first section, "Definitions," provides necessary definitions for working with pitch-class sets as bit strings. "Hierarchizability" begins by outlining the intuitive motivations underlying the idea of hierarchizability, proceeds to define hierarchizability more precisely, and finally shows how the universe of possible 12-pc set-classes breaks down into more and less hierarchizable set classes. "Hierarchizability of Scales in Western European, Northern Indian, and Japanese Traditional Art Music" explores how hierarchizability applies to scales common in Western European, Northern Indian, and Japanese traditional musics and finds that for these cultures, hierarchizability predicts scale candidacy ($p =$ .002). "Hierarchizability, Evenness, and Pattern Matching" revisits how hierarchizability relates to previous work on evenness and pattern matching in scale candidacy. "Limitations and Future Work" enumerates possible avenues for future research, including work on how hierarchizability relates both to cultures not looked at here and to extramusical domains like language. "Conclusion" summarizes the findings and ideas from the current investigation.

## Definitions

If $s$ is a length-$n$ bit string, $s$ is a vector containing $n$ values $s_0 \ldots s_{n-1}$, each of which is either `0` or `1`.

To allow for a common basis of comparison in a format that facilitates pattern matching, scales analyzed here are represented as bit string equivalents of 12-pc sets, where distances in 12-pc space are taken simply as representations of categorically equivalent intervals without implying any particular system of tuning or temperament. Given a length-12 sequence of 0's and 1's $s$ and a 12-pc set $S$, $s$ is a bit string representation of $S$ if for every $0 \leq i \leq 11$, $s_i = 1$ implies pitch-class $i$ is a member of $S$ and $s_i = 0$ implies pitch-class $i$ is not a member of $S$.

As with pitch-class sets, bit strings $a$ and $b$ can be related by transposition ($a_i = b_{(i+k) \bmod 12}$ for $0 \leq i \leq 11$), inversion ($a_i = b_{(k-i) \bmod 12}$), and complementation ($a_i = (b_i + 1) \bmod 2$).

Treating bit strings as equivalent under transposition, inversion, and complementation (i.e., as a way of encoding distance relationships, not information about absolute pitch) allows for an analogy between bit string classes and pc set-classes under which bit strings 101011010101, 110101011010, and 101001010 10 are equivalent representations of the diatonic and pentatonic set-classes 5-35 and 7-35. Unless noted otherwise, bit strings and pc sets are taken from here forward as members of bit string and pc set-classes respectively.

$|s|$ denotes bit string $s$'s length in bits.

Bit string $p$ is a substring of bit string $s$ ($p =\sim s$) if $|p| \leq |s|$ and there are some $i$, $k$ such that for every $0 \leq i < |p|$, $p_i = s_{(i+k) \bmod |s|}$.

Pc set-classes are identified using Forte numbers (Forte, 1973). Forte numbers group pitch collections by cardinality and assume equivalence under transposition and inversion but not complementation. For example, 7-35 denotes the 35$^{th}$ 7-pc set-class, which includes all diatonic modes.

## Hierarchizability

We begin with a thought experiment:

You are given a chunk of raw information in the form of the diatonic bit string 101011010101. However, you do not have any concept of the diatonic, or scale in general, or perhaps even music. This bit string represents pure structure—it does not need to be associated with any particular sensory modality. Furthermore, it is a different kind of information than a word, for example, in that it does not reference anything other than itself. Given these assumptions, then, your task is to figure out what this bit string *means*. What is the essence of this bit string, and how does its meaning differ from the meaning of other bit strings (keeping in mind that the meaning of any bit string is self-contained and does not make any external references whatsoever)?

This is a vague question demanding a precise answer. But intuitively, you already have some sense for what probably is *not* the meaning of this bit string. For instance, the fact that there are seven 1's does not really get at this bit string's unique essence since there are dozens of other possible length-12 bit strings containing seven 1's. The same can be said, for example, of the fact that there are no pairs of adjacent 0's. You could try interpreting the bit string in terms of an alphabet by segmenting it into $n$-bit units (e.g., letting $n = 2$ and treating 10, 11, etc. as unique "letters"), but the number of possible segmentations is limited only by the length of the bit string in question, suggesting this method does not really capture the bit string's fundamental "meaning."

To put it another way, these methods summarize certain local details of the bit string but do not capture these details in terms of their global context. Finding the unique "meaning" of this bit string entails understanding the bit string's high-level, global profile and being able to succinctly relate low-level, local details to this overarching structure—just as grasping the "meaning" of a sentence implies being able to relate individual word-level meanings to a unified sentence-level meaning.

Of course, whereas sentences consist of collections of rich semantic units tied together by relatively complex syntactic frameworks, bit strings are highly modular, apparently one-dimensional chains of extremely simple components. It would be reasonable to ask whether bit strings can really represent the kind of tightly integrated, globally interconnected structures sentences do—whether they can have any essential "meaning."

Bit strings do, however, contain high-level, holistic information in one sense: they contain large-scale repetitions of various substrings. Taking this observation as a starting point for interpreting the bit string whose "meaning" you are tasked with finding, then, you might notice that this "diatonic" bit string consists of a repetition of the substring 10101 with a 2-bit "remainder" tagged onto the end: (10101)(10101)01. This breakdown already provides some global context that can be used to situate individual bitwise components—for example, the initial 1 is no longer just the first element in a flat string, it is the first element of the first repetition of 10101 (the longest, or most prominent, non-overlapping repeating substring). Note that if we take this bit string to be circular by making its initial and final bits adjacent, the 2-bit "remainder" provides global context by differentiating substring repetitions: in the circular 12-bit string (10101)(10101)01 the two instances of 10101 can be uniquely identified as "the repetition preceding the remainder" and "the repetition following the remainder," but in the circular 10-bit string (10101)(10101) they can not be distinguished from each other in terms of the bit string's global structure.

It is possible to take this approach further, though, to find a more integrated, top-down "meaning" for the bit string 101011010101. Parsing this string in terms of repetitions of the flat substring 10101 reproduces the situation on a lower level: namely, although this breakdown in terms of the repeating substring 10101 gives each instance of this substring a unique position in the global context of the overall bit string, recursively breaking down each instance of 10101 into, for example, (10)(10)1 by latching onto a prominent second-level repeated substring gives rise to a more differentiated context within the first-level substring 10101.

This recursive breakdown pulls the integrated, multilayered structure ([10][10]1)([10][10]1)01 out of the superficially flat bit string 101011010101. In other words, one way to answer the question of this bit string's essential structure or meaning might be by arguing that 101011010101 = AAB where A = CCD, B = 01, C = 10, and D = 1. The "meaning" ([10][10]1)([10][10]1)01 is very well integrated in the sense that every local detail serves a unique function in terms of the bit string's global structure: for instance, the initial 1 is the 1 in the first repetition of the substring 10 in the repetition of the substring 10101 immediately following the 2-bit remainder in the circular string 101011010101. In other words, this hierarchical interpretation of bit strings tries to go after the gestalt of the matter by describing bit strings' unique global profiles in terms of nested, non-overlapping repeated substrings. This recursive method for reading integrated "meaning" into bit strings also fits with the analogy to language in that sentences similarly break down into constituent elements hierarchically rather than purely sequentially (see, for instance, syntax trees).

Now, we extend our preliminary thought experiment:

In light of your success determining the meaning of the bit string 101011010101, you are offered a contract under which you will be payed a large sum of money in exchange for finding the meanings of ten thousand additional bit strings. You accept the offer. As you work through the ten thousand bit strings whose meanings you are tasked with finding, breaking them down into hierarchies of repeating substrings, you notice there is considerable variety in the resulting structures.

The bit string 100111100111, for example, contains the prominent 6-bit repeated substring 100111, resulting in the first-level breakdown (100111)(100111). Unlike the diatonic 101011010101, then, 100111100111 contains no "position-finding" remainder to differentiate the two instances of the repeated substring 100111 given a circular version of the bit string—it is of the form AA instead of AAB.

On the other hand, the bit string 101000101011 breaks down into (1010)00(1010)11, a structure with a four-bit repeated substring and four non-adjacent remainder bits, on the first level. Each instance of the repeated substring 1010 is parsed as (10)(10), a structure with a two-bit repeated substring but no remainder bits, on the second level. Notice that (1) remainder bits on a given level undergo no recursive breakdown and share a single common function in the global context of the bit string and that (2) lower-level breakdowns that contain no remainder bits are less internally differentiated, as with analogous first-level breakdowns, so the relative abundance of first-level remainder bits and the lack of second-level remainder bits make the structure ([10][10])00([10][10])11 in a sense "flatter," less hierarchical, or less globally integrated than, for example, the diatonic ([10][10]1)([10][10]1)01.

The "augmented triad" bit string 100010001000 is yet another possible structural variation insofar as its first-level breakdown contains *three* instances of the prominent repeated substring 1000 (is of the form AAA). As with the earlier example (100111)(100111), a lack of "position-finding" bits in the structure (1000)(1000)(1000) mean individual instances of the first-level repeated substring 1000 are not uniquely identifiable in the circular version of the bit string, resulting in a less differentiated global context in which local details can be situated than in the case of the diatonic.

In general, these examples demonstrate that from the perspective of top-down parsing in terms of non-overlapping repeated substrings, not all bit strings are created equal. Differences in the resulting hierarchical structures mean some bit strings have more sharply defined global profiles or stronger, more integrated global identities that uniquely identify lower-level, local details in relation to the bit string as a whole. To push the analogy with language, some bit strings resemble complete sentences while others are akin to less unified sequences of grammatically recognizable fragments strung haphazardly together.

Given the range of possible structural variations and the gradation from highly integrated to more fragmentary types of "meaning" bit strings can be associated with, it would be useful to be able to quantify and compare different bit strings' capacity for assuming this kind of top-down structure—that is, their capacity for being organized through hierarchies of repeated substrings. This desire to categorize bit strings' ability to encode hierarchical structure in terms of prominent nested repeating patterns motivates the following definition of hierarchizability.

Moving beyond our thought experiment, although it is unlikely in reality that anyone would ever be interested in exchanging a large sum of money for the recursive breakdowns of ten-thousand bit strings, the problem of inferring the "meaning" of a bit string given little external guidance is one listeners (especially those with little previous musical exposure) confront every day. Specifically, listeners new to a given musical tradition are tasked with inferring scales' (more-or-less shared) meanings from the internal structures of the scales themselves. Idiomatic usage in actual music of course contributes to listeners' culture-specific understandings of basic musical building blocks, but picking up on this idiomatic usage in the first place entails already having some basic representation of the building blocks in question.

I hypothesize that parsing scales in terms of prominent, nested repeating scale fragments provides a powerful heuristic for quickly forming concise representations both of scales as wholes and of individual notes' shades of meaning in terms of their global scalar context; that scales whose bit string representations are more conducive to being broken down according to the process outlined above have stronger and more integrated global identities, are richer in potential meaning, and more easily facilitate implicit learning through listening; and that consequently, there is a culture-neutral preference for these more hierarchizable scales across world music.

### DEFINITION OF HIERARCHIZABILITY

Evaluating this last claim in particular requires a more precise definition of hierarchizability. $h_k(s)$, or hierarchizability with maximum remainder $k$ of bit string $s$, is defined recursively:

Let `max_repeat(s)` be any substring $p$ of $s$ such that there are $m \geq 2$ non-overlapping instances of $p$, $|s|$ - $m$ * $|p| \leq k$, there is no non-overlapping repeated substring $q$ of $s$ with $|q| > |p|$, and there is no non-overlapping repeated substring $r$ of $s$ where $|r| = |p|$ and making `max_repeat(s)` $r$ instead of $p$ ultimately results in a greater value for $h_k(s)$. Let `max_repeat(s)` = `NULL` if no such substring $p$ exists. Finally, let `remainder(s)` = $|s|$ - $m$ * |`max_repeat(s)`|.

If `max_repeat(s)` = `NULL` or if $s$ contains exclusively `1`'s or exclusively `0`'s, $h_k(s) = 1$.
If `remainder(s)` = 0,
$h_k(s) = h_k($`max_repeat(s)`$)$.
Otherwise, $h_k(s) = h_k($`max_repeat(s)`$) + 1$.

Within a given layer of recursion, $h_k(s)$ finds the longest repeated substring `max_repeat(s)` such that

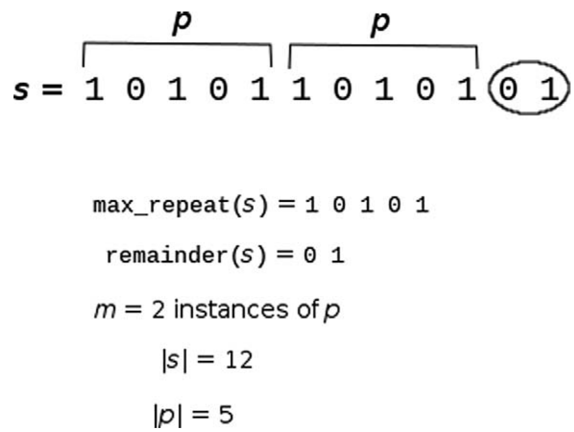


**FIGURE 1.** $h_k()$'s first-level parsing of the diatonic bit string $s$

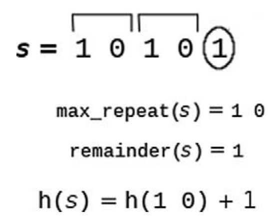


**FIGURE 2.** $h_k()$'s recursive evaluation of $s = 10101$

there are no more than $k$ remainder bits not belonging to some (non-overlapping) instance of `max_repeat(s)` (if there are two equal-length such subtrings, it chooses whichever one ultimately maximizes hierarchizability). If there is at least one remainder bit, $h_k(s)$ is $h_k($`max_repeat(s)`$) + 1$. If there are no remainder bits, individual instances of `max_repeat(s)` are not uniquely identifiable in terms of their global context, so $s$ is less hierarchizable and $h_k(s)$ is simply $h_k($`max_repeat(s)`$)$—individual bits and lower-level repeated substrings may still be differentiated within `max_repeat(s)`'s internal context, but they are not as globally integrated into $s$'s highest-level context as they would be if $s$ contained "position-finding" remainder bits at its highest level. Finally, in the base case, there is no repeated substring allowing for a $k$-bit ceiling on the remainder or $s$ contains only one kind of bit (making $s$ totally flat), so no further recursive breakdown is possible and $h_k(s)$ is 1.

Hierarchizability's $k$-bit cap on the remainder and preference for the *longest* repeating substring combine to establish a threshold of perceptual significance for `max_repeat(s)`. In other words, repeating substrings containing more bits are more noticeable, and strings with fewer (but more than zero) free-floating remainder

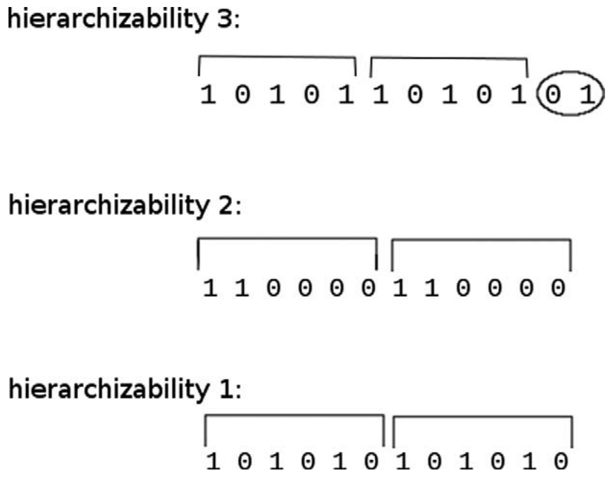


**FIGURE 3.** First-level breakdown of example strings with hierarchizability 1, 2, and 3 for $k = 2$

bits are more globally integrated, so $\mathtt{h}_k\mathtt{()}$ discards repeating substrings that have non-maximal lengths or that result in remainders exceeding the $k$-bit cutoff. For instance, the repeated substring 100 cannot really be said to determine the global structure of the string 100100111010 to the same extent, say, the substring 10101 does that of the diatonic 101011010101. Since breaking 100100111010 down into (100)(100)111010 leaves a 6-bit remainder, $\mathtt{h}_k\mathtt{(100100111010)} = 1$ for all $k < 6$.

The recursion's base case—i.e., the point where no further breakdown into repeated substrings occurs—also covers strings consisting of exclusively 1's or exclusively 0's. Such strings are simply assigned hierarchizability 1 to capture the intuition that, for example, 11111's breakdown into (11)(11)1 does not really reflect the string's ability to assume a top-down structure given a lack of true high-level position-finding bits (the "remainder" 1 is indistinguishable from the 1's in the "repeating substrings").

It turns out that for length-12 bit strings, hierarchizability can be as high as 3. Notice that for length-12 bit string $s$, max_repeat($s$) can be at most 5 for there to be any remainder bits in the first layer of recursion, making max_repeat(max_repeat($s$)) at most 2, leading to the base case in the third layer of recursion.

Letting $k = 2$, the diatonic provides an example of a scale with hierarchizability 3. $\mathtt{h}_2\mathtt{(101011010101)}$ proceeds according to the following steps:

1: max_repeat(101011010101) = 10101, leaving 2-bit remainder 01, so $\mathtt{h}_2\mathtt{(10101\ 10101\ 01)} = \mathtt{h}_2\mathtt{(10101)} + 1$

 2: max_repeat(10101) = 10, leaving 1-bit remainder 1, so $\mathtt{h}_2\mathtt{(10\ 10\ 1)} = \mathtt{h}_2\mathtt{(10)} + 1$

  3: max_repeat(10) = NULL, so $\mathtt{h}_2\mathtt{(10)} = 1$

 4: $\mathtt{h}_2\mathtt{(10101)} = 1 + 1 = 2$

5: $\mathtt{h}_2\mathtt{(101011010101)} = 2 + 1 = 3$

By contrast, the whole-tone 101010101010 contains no remainder bits at any level:

1: max_repeat(101010101010) = 101010, so $\mathtt{h}_2\mathtt{(101010\ 101010)} = \mathtt{h}_2\mathtt{(101010)}$

 2: max_repeat(101010) = 10, so $\mathtt{h}_2\mathtt{(10\ 10\ 10)} = \mathtt{h}_2\mathtt{(10)}$

  3: max_repeat(10) = NULL, so $\mathtt{h}_2\mathtt{(10)} = 1$

 4: $\mathtt{h}_2\mathtt{(101010)} = 1$

5: $\mathtt{h}_2\mathtt{(101010101010)} = 1$

Finally, the remainder-dominated scale mentioned earlier, 100100111010, undergoes a very simple "breakdown" for $k = 2$:

1: max_repeat(100100111010) = 100, leaving 6-bit remainder 111010, so $\mathtt{h}_2\mathtt{(100100111010)} = 1$

Since the parameter $k$ determines the upper limit on the number of remainder bits (i.e., the number of bits not included in an instance of the repeating substring), lower values of $k$ are more restrictive in the sense that they require the repeating substring more perceptibly dominate the structure of the bit string in question—so there are generally fewer maximally hierarchizable bit strings for smaller $k$. Table 1 shows hierarchizability under $k = 2$ and $k = 4$ for all possible length-12 bit string classes. The remainder of this study will focus on hierarchizability for these two values of $k$, following the reasoning that relaxing $k$ to 4 from 2 means significantly expanding the universe of possible breakdowns in terms of repeating substrings without dropping the requirement that a substantial majority of member bits participate in some instance of the repeating substring at any given level. Other values of $k$ are of course possible, but $k < 2$ precludes the possibility of any remainder bits whatsoever for length-12 strings, $k = 3$ represents a lower-contrast middleground between $k = 2$ and $k = 4$, and $k > 4$ results in a less interesting distribution of hierarchizabilites by virtue of its overinclusiveness.

Taking into account equivalence under transposition, inversion, and complementation, there are 122 unique bit

**TABLE 1.** *Hierarchizability for Length-12 Bit Strings*

| $s$ | Forte Number | $h_2(s)$ | $h_4(s)$ | $(h_2(s) + h_4(s))$ / 2 |
|---|---|---|---|---|
| 000000000000 | 0−1, 12−1 | 1 | 1 | 1 |
| 100000000000 | 1−1, 11−1 | 2 | 2 | 2 |
| 110000000000 | 2−1, 10−1 | 2 | 2 | 2 |
| 101000000000 | 2−2, 10−2 | 2 | 2 | 2 |
| 100100000000 | 2−3, 10−3 | 2 | 2 | 2 |
| 100010000000 | 2−4, 10−4 | 2 | **3** | 2.5 |
| 100001000000 | 2−5, 10−5 | **3** | **3** | **3** |
| 100000100000 | 2−6, 10−6 | 2 | 2 | 2 |
| 111000000000 | 3−1, 9−1 | 1 | 2 | 1.5 |
| 110100000000 | 3−2, 9−2 | 1 | 2 | 1.5 |
| 110010000000 | 3−3, 9−3 | 1 | 2 | 1.5 |
| 110001000000 | 3−4, 9−4 | 1 | **3** | 2 |
| 110000100000 | 3−5, 9−5 | **3** | **3** | **3** |
| 101010000000 | 3−6, 9−6 | 1 | 2 | 1.5 |
| 101001000000 | 3−7, 9−7 | 1 | **3** | 2 |
| 101000100000 | 3−8, 9−8 | **3** | **3** | **3** |
| 101000010000 | 3−9, 9−9 | **3** | **3** | **3** |
| 100100100000 | 3−10, 9−10 | **3** | **3** | **3** |
| 100100010000 | 3−11, 9−11 | **3** | **3** | **3** |
| 100010001000 | 3−12, 9−12 | 2 | 2 | 2 |
| 111100000000 | 4−1, 8−1 | 1 | 2 | 1.5 |
| 111010000000 | 4−2, 8−2 | 1 | 2 | 1.5 |
| 110110000000 | 4−3, 8−3 | 1 | 2 | 1.5 |
| 111001000000 | 4−4, 8−4 | 1 | 2 | 1.5 |
| 111000100000 | 4−5, 8−5 | 1 | **3** | 2 |
| 111000010000 | 4−6, 8−6 | **3** | **3** | **3** |
| 110011000000 | 4−7, 8−7 | 1 | 2 | 1.5 |
| 110001100000 | 4−8, 8−8 | **3** | **3** | **3** |
| 110000110000 | 4−9, 8−9 | 2 | 2 | 2 |
| 101101000000 | 4−10, 8−10 | 1 | 2 | 1.5 |
| 110101000000 | 4−11, 8−11 | 1 | 2 | 1.5 |
| 101100100000 | 4−12, 8−12 | 1 | **3** | 2 |
| 110100100000 | 4−13, 8−13 | 1 | 2 | 1.5 |
| 101100010000 | 4−14, 8−14 | **3** | **3** | **3** |
| 110010100000 | 4−Z15, 8−Z15 | 1 | 2 | 1.5 |
| 110001010000 | 4−16, 8−16 | 1 | **3** | 2 |
| 100110010000 | 4−17, 8−17 | **3** | **3** | **3** |
| 110010010000 | 4−18, 8−18 | 1 | **3** | 2 |
| 110010001000 | 4−19, 8−19 | 1 | **3** | 2 |
| 110001001000 | 4−20, 8−20 | **3** | **3** | **3** |
| 101010100000 | 4−21, 8−21 | 1 | 2 | 1.5 |
| 101010010000 | 4−22, 8−22 | 1 | **3** | 2 |
| 101001010000 | 4−23, 8−23 | **3** | **3** | **3** |
| 101010001000 | 4−24, 8−24 | 1 | **3** | 2 |
| 101000101000 | 4−25, 8−25 | 2 | 2 | 2 |
| 100101001000 | 4−26, 8−26 | **3** | **3** | **3** |
| 101001001000 | 4−27, 8−27 | 1 | **3** | 2 |
| 100100100100 | 4−28, 8−28 | 2 | 2 | 2 |
| 110100010000 | 4−Z29, 8−Z29 | 1 | **3** | 2 |
| 111110000000 | 5−1, 7−1 | 1 | 1 | 1 |
| 111101000000 | 5−2, 7−2 | 1 | 1 | 1 |
| 111011000000 | 5−3, 7−3 | 1 | 1 | 1 |
| 111100100000 | 5−4, 7−4 | 1 | 1 | 1 |
| 111100010000 | 5−5, 7−5 | 1 | **3** | 2 |
| 111001100000 | 5−6, 7−6 | 1 | 2 | 1.5 |
| 111000110000 | 5−7, 7−7 | **3** | **3** | **3** |

*(continued)*

**TABLE 1.** (continued)

| $s$ | Forte Number | $h_2(s)$ | $h_4(s)$ | $(h_2(s) + h_4(s))$ / 2 |
|---|---|---|---|---|
| 101110100000 | 5−8, 7−8 | 1 | 1 | 1 |
| 111010100000 | 5−9, 7−9 | 1 | 1 | 1 |
| 110110100000 | 5−10, 7−10 | 1 | 1 | 1 |
| 101110010000 | 5−11, 7−11 | 1 | **3** | 2 |
| 110101100000 | 5−Z12, 7−Z12 | 1 | 2 | 1.5 |
| 111010001000 | 5−13, 7−13 | 1 | **3** | 2 |
| 111001010000 | 5−14, 7−14 | 1 | 1 | 1 |
| 111000101000 | 5−15, 7−15 | **3** | **3** | **3** |
| 110110010000 | 5−16, 7−16 | 1 | 1 | 1 |
| 110110001000 | 5−Z17, 7−Z17 | 1 | **3** | 2 |
| 110011010000 | 5−Z18, 7−Z18 | 1 | 2 | 1.5 |
| 110100110000 | 5−19, 7−19 | **3** | **3** | **3** |
| 110100011000 | 5−20, 7−20 | **3** | **3** | **3** |
| 110011001000 | 5−21, 7−21 | 1 | 2 | 1.5 |
| 110010011000 | 5−22, 7−22 | **3** | **3** | **3** |
| 101101010000 | 5−23, 7−23 | 1 | 2 | 1.5 |
| 110101010000 | 5−24, 7−24 | 1 | 2 | 1.5 |
| 101101001000 | 5−25, 7−25 | 1 | **3** | 2 |
| 101011001000 | 5−26, 7−26 | 1 | **3** | 2 |
| 110101001000 | 5−27, 7−27 | 1 | 2 | 1.5 |
| 101100101000 | 5−28, 7−28 | **3** | **3** | **3** |
| 110100101000 | 5−29, 7−29 | **3** | **3** | **3** |
| 110010101000 | 5−30, 7−30 | 1 | 2 | 1.5 |
| 110100100100 | 5−31, 7−31 | **3** | **3** | **3** |
| 110010100100 | 5−32, 7−32 | **3** | **3** | **3** |
| 101010101000 | 5−33, 7−33 | **3** | **3** | **3** |
| 101010100100 | 5−34, 7−34 | 2 | 2 | 2 |
| 101010010100 | 5−35, 7−35 | **3** | **3** | **3** |
| 111010010000 | 5−Z36, 7−Z36 | 1 | 1 | 1 |
| 100111001000 | 5−Z37, 7−Z37 | 1 | **3** | 2 |
| 111001001000 | 5−Z38, 7−Z38 | 1 | **3** | 2 |
| 111111000000 | 6−1 | 1 | 1 | 1 |
| 111110100000 | 6−2 | 1 | 1 | 1 |
| 111101100000 | 6−Z3, 6−Z36 | 1 | 1 | 1 |
| 111011100000 | 6−Z4, 6−Z37 | 1 | **3** | 2 |
| 111100110000 | 6−5 | 1 | 2 | 1.5 |
| 111001110000 | 6−Z6, 6−Z38 | **3** | **3** | **3** |
| 111000111000 | 6−7 | 1 | 1 | 1 |
| 101111010000 | 6−8 | 1 | 1 | 1 |
| 111101010000 | 6−9 | 1 | 1 | 1 |
| 110111010000 | 6−Z10, 6−Z39 | 1 | **3** | 2 |
| 111011010000 | 6−Z11, 6−Z40 | 1 | 1 | 1 |
| 111010110000 | 6−Z12, 6−Z41 | 1 | 1 | 1 |
| 110110110000 | 6−Z13, 6−Z42 | 1 | **3** | 2 |
| 110111001000 | 6−14 | 1 | 1 | 1 |
| 111011001000 | 6−15 | 1 | 1 | 1 |
| 110011101000 | 6−16 | 1 | 2 | 1.5 |
| 111010011000 | 6−Z17, 6−Z43 | 1 | 2 | 1.5 |
| 111001011000 | 6−18 | 1 | 2 | 1.5 |
| 110110011000 | 6−Z19, 6−Z44 | **3** | **3** | **3** |
| 110011001100 | 6−20 | 1 | 1 | 1 |
| 101110101000 | 6−21 | 1 | 2 | 1.5 |
| 111010101000 | 6−22 | 1 | 2 | 1.5 |
| 101101101000 | 6−Z23, 6−Z45 | 1 | **3** | 2 |
| 110110101000 | 6−Z24, 6−Z46 | 1 | 2 | 1.5 |
| 110101101000 | 6−Z25, 6−Z47 | **3** | **3** | **3** |

*(continued)*

**TABLE 1.** (continued)

| $s$ | Forte Number | $h_2(s)$ | $h_4(s)$ | $(h_2(s) + h_4(s))$ / 2 |
|---|---|---|---|---|
| 110101011000 | 6−Z26, 6−Z48 | 1 | 2 | 1.5 |
| 110110100100 | 6−27 | 1 | 2 | 1.5 |
| 110101100100 | 6−Z28, 6−Z49 | 1 | 2 | 1.5 |
| 110100101100 | 6−Z29, 6−Z50 | 1 | 2 | 1.5 |
| 110100110100 | 6−30 | 2 | 2 | 2 |
| 110101001100 | 6−31 | 1 | 2 | 1.5 |
| 101011010100 | 6−32 | **3** | **3** | **3** |
| 101101010100 | 6−33 | 2 | 2 | 2 |
| 110101010100 | 6−34 | 2 | 2 | 2 |
| 101010101010 | 6−35 | 1 | 1 | 1 |

string classes, as shown in Table 1. For $k = 2$, 28 of these classes have hierarchizability 3, 13 2, and 81 1 while for $k = 4$, 52 have hierarchizability 3, 48 2, and 22 1. In general terms, a minority of bit strings are maximally hierarchizable under $k = 2$ while a significantly larger minority are maximally hierarchizable under $k = 4$. The predominance of hierarchizability-1 bit strings under $k = 2$ reflects $\mathtt{h}_2\mathtt{()}$'s strict limit on the number of possible remainder bits—most bit strings do not contain first-level repeated substrings whose instances encapsulate at least 10 of the 12 possible member bits.

A quick glance at at the 28 maximally hierarchizable bit strings under $k = 2$ is provocative. For instance, among 7-pc bit strings, 7-32 is the harmonic minor and 7-35 the diatonic. Among 6-pc bit strings, 6-Z47 is the blues hexatonic, 6-32 the Guidonian hexachord.

Using this partitioning of the universe of possible bit strings based on $\mathtt{h}_2\mathtt{()}$ and $\mathtt{h}_4\mathtt{()}$, it is now possible to evaluate more systematically hierarchizability as a culture-neutral predictor of scale candidacy.

## Hierarchizability of Scales in Western European, Northern Indian, and Japanese Traditional Art Music

Evaluating any property's claim to being a culture-independent contributor to scale candidacy means venturing into the murky area of making direct comparisons between pitch structures built on separate acoustic frameworks and finding analogies between different systems of tuning, temperament, and pitch classification that do not merely recast non-Western musical systems in terms of basic Western musical constructs. In order to avoid as much as possible the obvious potential for cultural bias working with pitch-class sets (specifically, bit-string equivalents of pitch-class sets) introduces, it is important to be clear about the assumptions hierarchizability makes in terms of tuning and temperament.

Critically, bit string representations of pitch collections and hierarchizability as applied to these bit string representations do not assume equal temperament, nor do $n$-bit strings assume an overarching pitch system akin to the Western 12-pc chromatic that breaks down into $n$ discrete pitches in actual musical practice. Instead, hierarchizability evaluates pitch collections in terms of internal pattern matching, and bit string representations as used by hierarchizability convey pitch collections' internal pattern matching profiles rather than any specific acoustic data. For the purposes of evaluating hierarchizability, a given bit string is a faithful representation of a scale taken from any musical system whatsoever as long as every pair of matching bit string segments corresponds to a pair of "matching" scale segments in actual practice and every pair of non-matching bit string segments corresponds to a pair of "non-matching" scale segments in actual practice.[1]

Since hierarchizability's essence lies first and foremost in its perceptual or cognitive significance, scale segments in actual musical practice are taken to be "matching" if they are *perceived* as matching by listeners familiar with the musical tradition in question—that is, if they are categorically equivalent.[2] Therefore, pattern matching as relevant to hierarchizability has no real acoustic meaning. Categorical equivalence, as the phenomenon that allows music to be parsed in terms of perceptual groupings rather than strict acoustic mandates, is the basis for pattern matching in hierarchizability and makes no demands in the area of tuning and temperament.

Of course, the assumption that pattern matching corresponds to categorical equivalence implies that a given bit string's claim to be a faithful representation of a certain pitch collection in a specific culture can be assessed only with the help of information on which intervals within that pitch collection are taken to be categorically equivalent within the relevant culture. Given the difficulties inherent in collecting these data, many musical cultures with scales whose conventional 12-pc translations are overrepresented among maximally hierarchizable bit strings are mostly excluded from the present study. See, for example, Javanese and Balinese gamelan: the closest 12-pc approximations of the 5-tone *pelog*

[1] To see why this independence from equal temperament is helpful in working with scales from world music, consider that some cultures use intervals smaller than the semitone without implying any kind of 24-pc chromatic breakdown (Burns, 1999, p. 217).

[2] See Patel (2008, p. 22) for an overview of categorical perception in music.

and 7-tone *pelog* scales (5-20 and 7-29 respectively) are maximally hierarchizable, but the basic cognitive underpinnings of these scales and the tuning systems in which they are situated remain so poorly understood from an analytical perspective that it is impossible to say with any real certainty whether these pitch sets accurately convey the perceptual structure (in terms of categorical equivalence) of the scales they allegedly represent.

For this reason, the present study is restricted to the central scales of three practices whose pitch systems allow for relatively straightforward conversion to 12-pc sets and length-12 bit strings: Western European art music, Northern Indian classical music, and Japanese traditional music. Sidestepping the issues posed by using pitch systems with more involved, controversial conversions allows for the assumption that traditional interval classes in 12-pc space correspond to classes of categorically equivalent intervals in actual practice. Specifically, the scales taken here as representative of these cultures are: the diatonic (7-35), harmonic minor (7-32), and ascending melodic minor (7-34) in Western European music; Bhatkhande's ten *thāts* in Northern Indian music (six instances of 7-35, two instances of 7-20, one instance or 7-22, and one instance of 7-29); and the *in* (5-20) and *yo* (5-29) families of scales in Japanese music.

This combined list of scales, which includes six unique pitch-class sets (5-20 / 7-20, 5-32 / 7-32, 7-22, 7-29, 7-32, 9-7) after taking into account complementary equivalence and cross-cultural overlap, represents an attempt to distill a collection of core scales out of these complex and nuanced musical traditions. In the case of Western European art music, the three primary scales of common-practice music seem like an obvious choice. For Northern Indian classical music, selecting a smaller set of core scales means generalizing a set of parent modes out of the multitude of *ragas* common in musical practice. Although theorists like Jairazbhoy (1971, p. 181) have put forward more comprehensive alternatives to Bhatkhande's traditional system of categorizing ragas in terms of ten *thāts*, Jairazbhoy (1971, p. 54) points out that "Bhatkhande's ten *thāts* give an indication of the principal scales used in present-day North Indian classical music—scales which have withstood the test of time." Since the goal here is to explore the hypothesis that more hierarchizable pitch collections are more likely to appear as scales in world music, focusing on the conventional, less expansive ten-*thāt* framework excludes marginal or "constructed" scales with limited significance in Northern Indian musical practice. Along similar lines, the present study differs from Carey (2002) and other previous work that has relied on the much more theoretically exhaustive system of seventy-two *melakarta* common in Southern India. This choice is motivated by the *melakarta*'s inclusion of more "experimental" scales that occur rarely in practice and have fallen into use primarily as a result of their appearance in the *melakarta* in the first place (Jairazbhoy, 1971, p. 48).

Finally, the *in* and *yo* families of scales are included as representative of the most prominent scales in traditional Japanese music. The corresponding pitch collections, 5-20 and 5-35 respectively, encompass the four typical scales built on like "tetrachords" that Koizumi (1977) lists—*minyo onkai*, *miyakobushi onkai*, *ritsu onkai*, and *ryukyu onkai*. The more rarely heard family of scales built on unlike tetrachords (5-28) is omitted.

Beyond their natural connection to the language of 12-pc sets, Western European, Northern Indian, and Japanese traditional musics provide especially fruitful material for the current study for two other reasons. First, these cultures provide a comparatively broad and varied pool of scales to draw on. Second, all three musics are historically distinct from one another, making it easier to assess hierarchizability as a culture-independent contributor to scale candidacy—any shared scales are not simply the product of a common musical history.

That said, there is no way to rule out the possibility that focusing on pitch systems with more direct translations to 12-pc sets introduces some cultural bias. It could be the case that musical systems that break down the octave into 12 discrete pitches share other common properties and that focusing on such systems means evaluating hierarchizability not as a truly culture-neutral predictor of scale-candidacy but as a predictor of scale candidacy in certain types of musical traditions. That 12-pc approximations of scales fundamental to musics with less familiar tuning systems repeatedly appear among optimally hierarchizable pitch collections—see the Balinese 5- and 7- tone *pelog* (5-20, 7-29), the Arabic *nahawand* aggregate (8-26), or the Turkish *puselik* (7-32), for example—suggests these results may generalize to cultures that do not divide the octave into twelve discrete pitches, but further work is necessary to systematically clarify both bit strings' ability to encode patterns based on categorically equivalent intervals in less 12-pc-like pitch systems and hierarchizability's significance for non-length-12 bit strings.

Table 2 lists the basic scales from Western European, Northern Indian, and Japanese traditional music described above along with Forte number, bit string representation $s$, and hierarchizability for $k = 2$ and $k = 4$. Examples of valid recursive breakdowns for $k = 2$ are also given.

Of the Western European scales in Table 2, the major, natural minor, and harmonic minor are all maximally hierarchizable for both $k = 2$ and $k = 4$. Notice that the harmonic minor bit string can be thought of as a copy of the natural minor bit string with the remainder bits flipped (`1011010110`**`10`** → `1011010110`**`01`**).

The ascending melodic minor is the only Western European scale—and indeed, the only scale—on the list that is not optimally hierarchizable for all values of $k$. Specifically, pc set-class 7-34 has hierarchizability 2 for both $k = 2$ and $k = 4$. Although the ascending set is taken here as representative of the melodic minor, using the aggregate set 9-7 (in analogy to *thāts*, which function as *raga* "supersets") does not materially change the results as 9-7 has hierarchizability 1 for $k = 2$ and 3 for $k = 4$. Both bit strings have an average hierarchizability of 2 over $k = 2$ and $k = 4$, and from either perspective, the melodic minor is more hierarchizable than average but not optimally hierarchizable. This result is at odds with Carey's (2002) coherence quotient, under which the melodic minor 7-34 is second only to the diatonic 7-35 among 7-pc sets. The hypothesis that more hierarchizable bit strings are more likely to correspond to scales actually used in world music would suggest that given its lower status in terms of hierarchizability, 7-34 is less likely to appear as a scale in multiple independent musical traditions than, say, 7-32—and indeed, harmonic minor-like scales are abundant in world music (see, for instance, the previously mentioned Turkish *puselik* and Arabic *nahawand*) while melodic minor-like scales are, to the best of my knowledge, scarce.

One possible explanation for Western European art music's blend of more and less hierarchizable scales could be that unlike optimally hierarchizable scales, which are more universally preferred, suboptimally hierarchizable scales tend to arise only because they exhibit desirable properties in the context of certain cultures. For instance, while the melodic minor is unremarkable in and of itself from hierarchizability's perspective, it is especially useful as an outgrowth of the diatonic set given the specific practices of Western tonal harmony. This interpretation might suggest that maximally hierarchizable bit strings for $k = 2$ are more likely to emerge as scales independently across multiple cultures whereas maximally hierarchizable sets for $k = 4$ and intermediately hierarchizable sets for $k = 2$ are simply more likely to emerge as scales in particular cultures under additional, culture-specific influences (or, in the case of the melodic minor, as extensions of sets that are maximally hierarchizable for $k = 2$).

In any event, the melodic minor is the only suboptimally hierarchizable scale listed in Table 2. All ten Northern Indian *thāts* have hierarchizability 3 for $k = 2$. Six of the *thāts* are rotations of the diatonic and two of the remaining *thāts* are related by inversion, making for four pc sets altogether: 7-35, 7-20, 7-22, and 7-29. Like the diatonic 7-35, the latter three have precedent elsewhere in world music: 7-20's complement, 5-20, corresponds to the Japanese *in* as well as the closest 12-pc approximation of the Balinese/Javanese five-tone *pelog*; meanwhile, 7-22 corresponds to the closest 12-pc approximation of the Arabic *hijaz kar* and *nawa athar* while 7-29 corresponds to that of the Balinese/Javanese seven-tone *pelog*. Although the precise relationship between these length-12 bit strings and the scales they approximate remains unclear, this overlap with the Northern Indian *thāts* suggests that future work on the relevance of hierarchizability and (length-12) bit strings to these traditions could be revealing.

Similarly, the bit strings representing the Japanese *in* and *yo* families of scales both have hierarchizability 3 for $k = 2$. Koizumi (1977) influentially analyzed scales in Japanese music in terms of the constituent "tetrachords" [0, 3, 5], [0, 1, 5], [0, 2, 5], and [0, 4, 5]—for instance, *minyo onkai* in the *yo* family would consist of stacked, overlapping copies of [0, 3, 5], resulting in [0, 3, 5, 7, 10]. Under this system, scales in the common *in* and *yo* families are built out of stacked like tetrachords. Scales built out of unlike tetrachords and corresponding to pc set 5-28, though rarer (and therefore not included in Table 2), do appear in traditional Japanese music and have hierarchizability 3 for $k = 2$.

Thus, of the 10 scale categories in Table 2, 9 are maximally hierarchizable for $k = 2$. Taking into account equivalence under transposition, inversion, and complementation, these 10 scale categories draw on 6 unique bit strings (5-20 / 7-20, 7-22, 7-29, 7-32, 7-34, and 5-35 / 7-35), of which 5 have hierarchizability 3 for $k = 2$. Given that there are 28 bit strings with hierarchizability 3 for $k = 2$ out of 122 total, the probability of arriving at a result at least this strong under the null hypothesis (that more hierarchizable bit strings are not more likely to appear as scales across world music) is:

$$\left(\frac{28}{122}\right)^6 + \frac{\binom{28}{5}\binom{94}{1}}{\binom{122}{6}} = .002$$

For $k = 4$, 9 of the 10 scale categories once again have hierarchizability 3. Since 52 of 122 possible unique bit strings have hierarchizability 3 under $k = 4$, the probability of obtaining this result by chance is:

**TABLE 2.** *Hierarchizability for Western European, Northern Indian, and Japanese Traditional Scales*

| Scale | Forte Number | $s$ | $h_2(s)$ | $h_4(s)$ | Example Hierarchization |
|---|---|---|---|---|---|
| Western European major, natural minor | 7−35 | 101011010101 | 3 | 3 | ([10][10]1)([10][10]1)01, $k = 2$ |
| Western European harmonic minor | 7−32 | 101101011001 | 3 | 3 | ([10]1[10])([10]1[10])01, $k = 2$ |
| Western European ascending melodic minor | 7−34 | 101101010101 | 2 | 2 | (10)1(10)(10)(10)(10)1, $k = 2$ |
| Northern Indian *kalyan*, *bilaval*, *khamaj*, *kafi*, *asavri*, *bhairvi* | 7−35 | 101011010101 | 3 | 3 | ([10][10]1)([10][10]1)01, $k = 2$ |
| Northern Indian *tori* | 7−20 | 011101001110 | 3 | 3 | (0[1][1][1]0)10(0[1][1][1]0), $k = 2$ |
| Northern Indian *purvi* | 7−20 | 011100101110 | 3 | 3 | (0[1][1][1]0)01(0[1][1][1]0), $k = 2$ |
| Northern Indian *bhairav* | 7−22 | 110011011001 | 3 | 3 | ([1][1]00[1])10([1][1]00[1]), $k = 2$ |
| Northern Indian *marva* | 7−29 | 100101101011 | 3 | 3 | 10([01][01]1)([01][01]1), $k = 2$ |
| Japanese *in* (*miyakobushi*, *ryukyu*) | 5−20 | 101100011000 | 3 | 3 | 10(11[0][0][0])(11[0][0][0]), $k = 2$ |
| Japanese *yo* (*minyo*, *ritsu*) | 5−35 | 101010010100 | 3 | 3 | 10([10][10]0)([10][10]0), $k = 2$ |

$$\left(\frac{52}{122}\right)^6 + \frac{\binom{52}{5}\binom{70}{1}}{\binom{122}{6}} = .051$$

Furthermore, this correlation between hierarchizability and scale candidacy in actual musical practice is not merely attributable to hierarchizability's preference for 5- and 7-note pc sets: the probability that at least five of six randomly selected unique 5- and 7-pc set classes would have hierarchizability 3 for $k = 2$ is .005. It appears, therefore, that hierarchizability really does predict scale candidacy across multiple independent musical traditions, and I suggest that this association between hierarchizability and scale candidacy is not obviously reducible to some third variable.

## Hierarchizability, Evenness, and Pattern Matching

There is a fair amount of overlap between hierarchizability's categorization of pitch collections and those of previously explored set-theoretic properties—that is, sets that are maximally hierarchizable tend to display other exceptional properties and vice-versa. For instance, the five optimally hierarchizable 7-pc collections included in Table 2 all appear among the ten most coherent pc set-classes under Carey's (2002) ordering of 7-pc set-classes by coherence quotient.

Nevertheless, hierarchizability does differ on some accounts from previously explored set-theoretic properties that it shares common ground with. Perhaps most obviously, T-symmetrical sets are, as a rule, not maximally hierarchizable. By contrast, collections like the whole-tone have coherence quotients of 1, higher than those of any of the non-diatonic/pentatonic scales in Table 2, according to Carey's model. Similarly, the whole-tone collection is maximally even, whereas most of the scales listed in Table 2 are not (Clough & Douthett, 1991).[3]

Despite this discrepancy between hierarchizability and its theoretical forebears, however, actual musical practice is unambiguous on the topic of T-symmetrical scales: across world music, T-symmetrical scales are relatively rare. Scales from world music that are sometimes construed as T-symmetrical in the literature tend in actuality either to not be as equally spaced as first glance might suggest (see the Balinese/Javanese *slendro*; for example, Herbst, 1997, p. 33) or to function not as scales per se but as 12-pc-like fields from which "diatonic" subsets are drawn (see the Thai heptatonic; Morton, 1976, p. 126).

Even in Western music, T-symmetrical scales occupy a rather circumscribed space and act more often as local "effects" than true scales. Composers have long known that T-symmetrical collections are often most useful as extensions of more hierarchically defined scales—for instance, Persichetti (1961, p. 54) notes that "the true value of the whole-tone scale lies in the contrast it provides when it is used in combination with other scales and techniques." Twelve-tone serialism provides maybe the best example of music genuinely rooted in a T-symmetrical scale although it is certainly up for debate whether dodecaphony is a useful data point from

[3] Many of the scales in Table 2 also fail to exhibit well-formedness (Carey & Clampitt, 1989) as well as properties that distinguish the diatonic from the whole-tone like Myhill's property and structure implies multiplicity (Clough & Myerson, 1985).

the perspective of exploring broader, universal constraints on music.

On the theoretical front, the privileged status of T-symmetrical scales results not from an attempt to model some far-reaching significance of T-symmetrical scales in actual musical practice but rather from an extrapolation of the diatonic's exceptional evenness in the context of 7-pc sets into the question of scale candidacy among *n*-pc sets. This extrapolation is a key point of difference with hierarchizability: rather than attempting to generalize traits of particular 5- and 7-pc sets vis-a-vis other 5- and 7-pc sets to *n*-pc sets, hierarchizability takes the universe of all possible scales as its point of departure and attempts to home in on *n*-pc scale candidates from there. The contrast between these two approaches is evident in the fact that hierarchizability under $k = 2$ is inherently strongly biased towards 5- and 7-pc sets over 6-pc sets (29% of 5- / 7-pc sets vs. 11% of 6-pc sets have hierarchizability 3 for $k = 2$) whereas coherence quotient and maximal evenness require adding on set size as a second variable to distinguish the diatonic from the whole-tone.

Ultimately, the tension of the similarities and differences between hierarchizability and these other properties comes down to the concept of *evenness*. Evenness is central to coherence quotient and, of course, maximal evenness, each of which extrapolates the diatonic's outstanding evenness in relation to other 7-pc collections across *n*-pc space. Moreover, since sets that are richer in nested repeating stepwise patterns also tend to be more even, optimally hierarchizable sets are typically more even than average—hence the correlation between hierarchizability and coherence quotient, maximal evenness, etc.

However, evenness is not at the core of hierarchizability. Maximally hierarchizable sets may tend to be more even, but hierarchizability is not essentially another measure of evenness, which is where it largely differs from previous models of scale candidacy. Whereas the most even pc sets possible—T-symmetrical sets—appear as ideal scale candidates under evenness-based models (without introducing additional parameters that restrict, for instance, "diatonic" cardinality), these highly even sets are not seen as especially promising scale candidates from the perspective of hierarchizability.

One explanation for evenness-based models' preference for T-symmetrical scales is that in actual musical practice, scale candidacy is a function of both "evenness" and "interestingness" such that the diatonic is both "even" and "interesting" while the whole-tone is "even" but not "interesting." The question then becomes how to define both "evenness" and "interestingness." Relating this problem back to hierarchizability, there are at least two possible explanations for the divergence between hierarchizability and evenness-based models: either (1) hierarchizability captures aspects of both "evenness" and "interestingness" whereas coherence quotient and friends capture aspects of "evenness" but not "interestingness or (2) "evenness" is actually not particularly central to the question of scale candidacy but is rather a side effect of some other variable that determines scale candidacy, and hierarchizability captures aspects of this other variable. Further work should be able to begin unraveling the cause-and-effect at play here.

At its core, hierarchizability has more in common with properties based on "pattern matching" than those based on evenness. Most notably, hierarchizability applies Browne's pattern matching and position finding to stepwise scale fragments rather than intervals. Instead of being organized in terms of hierarchies of more and less common *intervals*, maximally hierarchizable sets are organized in terms of more and less common *scale fragments*. By definition, a given pitch collection contains more than one instance of any scale fragment that belongs to a repeated substring, and scale fragments belonging to lower-level repeated substrings (e.g., `10` in the harmonic minor `([10]1[10])([10]1[10])01`) appear more frequently than those belonging only to higher-level repeated substrings. Meanwhile, bits not belonging to repeated substrings allow for more distinctive "position finding" scale fragments—exactly what T-symmetrical sets lack.

Browne's notions of "interval content" and "interval context" are also relevant, again allowing stepwise scale fragments to replace intervals as the basic units of pattern matching. In particular, if *scale fragment content* refers to a given scale fragment's bit-by-bit content and *scale fragment context* refers to the bit-by-bit content of the portion of the total scale not belonging to a given scale fragment, any two instances of a given repeated substring `max_repeat(`*s*`)` within a maximally hierarchizable (for $k = 2$) 12-bit string *s* are identical in terms of scale fragment content but differ in terms of scale fragment context.

For example, the first level of the harmonic minor `(10110)(10110)01` contains two instances of `10110` that are uniquely positioned in relation to the two "remainder" bits and have unique overall scale fragment contexts (specifically, the scale fragment context of the first instance is . . . . . `1011001` while that of the second is `10110` . . . . . `01`—note that direction matters for scale fragment content/context but not interval content/context). More generally, unique instances of max_repeat(*s*) (or max_repeat(max_repeat(*s*))) have unique scale fragment contexts because if *s* contained two matching non-overlapping substrings that were identical in terms of both scale fragment content and scale fragment context, *s* would by definition be T-symmetrical, which would contradict the assumption that *s* is a maximally hierarchizable length-12 bit string.

Thus, saying that more hierarchizable strings are more tightly packed with uniquely identifiable non-overlapping nested repeating substrings is equivalent to saying that more hierarchizable strings are richer in non-overlapping nested substrings that are identical in terms of content but not context.

This interplay between content and context links hierarchizability to previous work establishing the importance of hierarchy in pitch perception. For example, Deutsch (1980) showed that listeners are better able to remember segmented pitch sequences with straightforward hierarchical encodings—in other words, pitch sequences that break down into nodes that are identical in content but different in context. Whereas individual instances of max_repeat(*s*) for pitch collection *s* existing in pitch-class space can only be uniquely differentiated with the help of position-finding bits, temporal position fulfills the position-finding function for short pitch sequences existing in time.[4]

By looking at hierarchizability in terms of scale fragment content and scale fragment context, it is also possible to see why scales that are related by inversion and complementation are necessarily equally hierarchizable: for a given string, neither swapping 1's and 0's nor reversing the order of the bits changes the fact that individual instances of max_repeat(*s*) are identical in content but unique in context.

Similarly, stepwise pattern matching and position finding point to hierarchizability's perceptual significance. When the pentatonic 101001010100 breaks down into ([10][10]0)([10][10]0)00, the component [10][10]0 has meaning only as a matched pattern within the scale as a whole. It is a hierarchical node in the scale's structural breakdown in terms of repeating scale fragments, not an isolated unit with independent musical significance. Treating the 1's as sounded pitches and the 0's as "empty space," the role of the 0's is that they are not 1's: flipping 0's to 1's or eliminating 0's by shortening the string to 11 bits would result in bit strings like 101101010100 or 10101010100, which no longer have the pentatonic's distinctive grouping in terms of repeating scale fragments. Of course, one could just as easily say that the role of 1's is that they are not 0's, which is why complementary bit strings are inevitably equally hierarchizable.

[4] As far as pitch, the fact that these sequences are often based on subsets of non-T-symmetrical scales introduces a further element of position finding.

## Limitations and Future Work

It is clear that in Western European, Northern Indian, and Japanese traditional musics, maximally hierarchizable sets occur as scales at a rate far beyond what would be expected based on chance. Nevertheless, further work is needed to pin down more exactly hierarchizability's significance as a universal predictor of scale candidacy.

Since the current study focused on musical traditions whose scales have straightforward 12-pc equivalents, future work should look at how hierarchizability applies to scales with more elusive bit string representations—for instance, gamelan scales and Arabic *maqamat*. It is promising that 12-pc approximations of the gamelan 5- and 7-tone *pelog* have hierarchizability 3 and that familiar diatonic and harmonic minor-like scales appear among Arabic *maqamat*, but only studies looking specifically at how listeners perceive intervals within these scales will be able to reliably resolve the question of whether accurate bit string representations of these scales exist.

Recall that from the perspective of hierarchizability, bit strings are only true representations of specific scales insofar as pairs of equally spaced bits correspond to endpoints of categorically equivalent intervals and more distantly spaced bits correspond to intervals perceived as categorically "greater" in actual musical practice. For example, 100000010000 does not denote all instances of "the perfect fifth" but only those situated in a pitch universe whose system of categorical relationships suggests this 12-bit breakdown. Therefore, one way to clarify the relationship between bit string representations and particular scales from world music would be to conduct perception studies leading to mappings of categorically equivalent intervals within these scales.

There is also the question of how hierarchizability applies to *n*-pc sets where $n \neq 12$. It seems probable that scales from some non-Western cultures will be best represented as non-length-12 bit strings although the apparent bias towards 12-pc sets evident in traditions like Northern Indian and Japanese music is in itself interesting (assuming it does not merely reflect a bias on the part of Western theorists towards exploring musics based on 12-pc systems).

Possibly the most interesting area for future research, however, has to do with how hierarchizability hints at parallels between music and language. At its essence, hierarchizability measures the extent to which bit strings (scales) have stronger or weaker and more or less integrated global identities—hierarchizability is fundamentally different than properties like coherence quotient in that it can only be calculated in a top-down, global, recursive way as opposed to in an incremental, interval-by-interval fashion.

This issue of how structures formed out of basic constituent units acquire coherent, global identities is as relevant to language as it is to music—to put it another way, both language and music are very hierarchical systems. In semantics, the question can be phrased in terms of how individual words converge into a unified meaning. In syntax, it can be thought of in terms of how parts of speech combine to form grammatically complete sentences.

Even in language acquisition, this question repeats itself on a basic auditory processing level in the problem of text segmentation: given a one-dimensional string of sound, how do you organize what you are hearing into groups of syllables? Cowan (1991) has approached this matter from the perspective of "the repetition of fixed sequences of speech"—a sort of syllabic analog to repeated substrings in the context of musical scale bit string equivalents. In Cowan's study, learned syllable groupings within a longer stream of speech were determined by "the size of the embedded segment, the repetition frequencies of the two pre-exposed patterns, and the serial position of each pre-exposure." This result suggests that the auditory system may generally gravitate to more prominent repeated substrings as a way of extracting meaning from flat structures and that a universal preference for more hierarchizable scales may reflect a more general attentive bias towards auditory structures richer in (positionally differentiable) repeating substrings.

Since strings of text exist in time and "strings" of pitches exist in pitch space, this parallel has implications in terms of Pressing's (1983) proposed isomorphisms between time and pitch space. Additionally, because "pattern matching" acts as a tool for learning and language acquisition in the context of text segmentation, Cowan's (1991) study raises the possibility that nested repeating substrings are especially salient in music from the perspective of "scale acquisition" and that more hierarchizable scales are also easier to learn.

Note that in text segmentation—as in pitch class hierarchization within musical scales—the task of inferring a given structure's integrated "meaning" is inextricably bound up with the task of grouping that structure's constituent elements. Indeed, text "segmentation" is fundamentally a global, contextual process—recent research has shown, for instance, that distal (in addition to proximal) prosody affects segmentation of familiar words and influences learning of novel words in ambiguous syllable sequences and that this global "position finding" function of prosody has observable neurological correlates (Breen, Dilley, McAuley, & Sanders, 2014; Dilley & McAuley, 2008).

Also along the lines of isomorphisms between time and pitch space, hierarchizability relates to work that has been done on tiling in music—much of which has concentrated on rhythmic tiling (Rahn & Amiot, 2011). Specifically, hierarchizability can be thought of as a way of measuring recursive "tileability" in pitch-class space, allowing for "position finding" or "remainder" tiles. Future work should explore whether this connection opens up any new mathematical insights into universal constraints on scales or parallels between tessellatation in time and pitch space. Tiling may point to further parallels with language too, since tasks like text segmentation and syntax parsing can be thought of as processes that involve recursively covering a given sentence "space" with predefined word or part-of-speech "tiles."

Finally, in light of the fact that the universe of optimally hierarchizable bit strings includes representations of structurally important pitch sets not conventionally thought of as "scales" like the major/minor triad and perfect fifth, future work should clarify the deceptively simple question of what exactly a scale is. Can these other pitch sets be reasonably called scales, is hierarchizability relevant more broadly to structurally important pitch collections rather than specific to scales, or is the appearance of these "non-scalar" pitch sets mere coincidence?

Previous work suggests that maximally hierarchizable lower-cardinality pitch collections like the major/minor triad may have some perceptually desirable properties compared with their less hierarchizable counterparts. Trainor and Trehub (1993) found that both infants and adults were better at detecting small changes to melodies based on major than augmented triads and perfect than augmented fifths.

This question as to the significance of maximally hierarchizable pitch collections with cardinality less than five highlights a limitation of the current study: while it is clear that optimally hierarchizable scales are overrepresented in Western European, Northern Indian, and Japanese traditional musics, the role of maximally hierarchizable pitch collections that do not arise in these traditions is more ambiguous. There are three main ways of accounting for these collections. First, they could represent scales used in world music that do not appear in the three traditions considered here (a possibility hinted at by the appearance of collections like 8-26, the Arabic *nahawand* aggregate). Alternatively, they could represent musical objects other than scales (a possibility hinted at by the appearance of the major/minor triad). Finally, they could represent pitch collections that are not used in world music because despite being maximally hierarchizable, they fail on some other cognitive, acoustic, or historical grounds. It is certainly possible that some mixture of all three factors is at play. Studies incorporating scales from cultures not considered here are likely to provide further insight.

## Conclusion

Throughout world music, scales are used to encode information about pitch hierarchy. With this common musical function of scales in mind, the current study has explored the hypothesis that some collections are more able to hold hierarchically organized pitch information and that these collections are more likely to appear as scales in actual musical practice.

A scale's hierarchizability ($\mathrm{h}_k(s)$ for bit string $s$) recursively measures the degree to which that scale is saturated with uniquely positioned repeating scale fragments (substrings). More hierarchizable scales (for smaller $k$) seem to have stronger top-down, global identities, making them better containers for learned, culture-specific hierarchical relationships.

At any rate, maximally hierarchizable pitch collections appear to be overrepresented among scales in world music. Five of the six pc set-classes that account for the most common scales in Western European, Northern Indian, and Japanese traditional music have hierarchizability 3 for $k = 2$. The likelihood of such a high portion of these scales being maximally hierarchizable under the null hypothesis is .002, suggesting that hierarchizability is a culture-independent predictor of scale candidacy.

The distribution of maximally hierarchizable scales is also consistent with a bias towards 5- and 7-pc scales over 6-pc scales in musical practice. Nevertheless, there are some maximally hierarchizable 6-pc sets, including equivalents of the Guidonian hexachord and blues hexatonic.

Nothing about $\mathrm{h}()$ is specific to equal temperament or to the 12-pc universe – or, ultimately, even to pitch space. This generality inherent in hierarchizability highlights the potential for future work dealing with a wider array of scales and tuning systems, cognitive isomorphisms between pitch and time, and parallels between music and language.

Perhaps most importantly, hierarchizability is inherently global and contextual in its conception of scales—"pattern matching" between repetitions of substrings and "position finding" in terms of non-repeated scale fragments are both non-incremental, gestalt processes. It may be that grasping the underlying cognitive constraints on musical objects like scales and the extramusical connections between these objects and domains like language requires going beyond additive properties rooted in local operations on independent elements to look at top-down properties that can only be observed by treating musical objects as interconnected, unified wholes.

## Author Note

*Correspondence concerning this article should be addressed to* Niels J. Verosky. E-mail: niels.verosky@gmail.com

## References

Breen, M., Dilley, L. C., McAuley, J. D., & Sanders, L. D. (2014). Auditory evoked potentials reveal early perceptual effects of distal prosody on speech segmentation. *Language, Cognition, and Neuroscience*, *29*, 1132-1146.

Browne, R. (1981). Tonal implications of the diatonic set. *In Theory Only*, *5*(6-7), 3-21.

Burns, E. M. (1999). Intervals, scales, and tuning. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 215-264). San Diego, CA: Academic Press.

Carey, N. (2002). On coherence and sameness, and the evaluation of scale candidacy claims. *Journal of Music Theory*, *46*, 1-56.

Carey, N., & Clampitt, D. (1989). Aspects of well-formed scales. *Music Theory Spectrum*, *11*, 187-206.

Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music of North India. *Journal of Experimental Psychology: General*, *113*, 394-412.

Clough, J., & Douthett, J. (1991). Maximally even sets. *Journal of Music Theory*, *35*, 93-173.

Clough, J., & Myerson, G. (1985). Variety and multiplicity in diatonic systems. *Journal of Music Theory*, *29*, 249-270.

Cowan, N. (1991). Recurrent speech patterns as cues to the segmentation of multisyllabic sequences. *Acta Psychologica*, *77*, 121-135.

DEUTSCH, D. (1980). The processing of structured and unstructured tonal sequences. *Perception and Psychophysics*, *28*, 381-389.

DILLEY, L. C., & MCAULEY, J. D. (2008). Distal prosodic context affects word segmentation and lexical processing. *Journal of Memory and Language*, *59*, 294-311.

FORTE, A. (1973). *The structure of atonal music*. New Haven, CT: Yale University Press.

HERBST, E. (1997). *Voices in Bali: Energies and perceptions in vocal music and dance theater*. Hanover NH: Wesleyan University Press.

HERHOLZ, S. C., LAPPE, C., & PANTEV, C. (2009). Looking for a pattern: An MEG study on the abstract mismatch negativity in musicians and nonmusicians. *BMC Neuroscience*, 10.

HUETTEL, S. A., MACK, P. B., & MCCARTHY, G. (2002). Perceiving patterns in random series: Dynamic processing of sequence in prefrontal cortex. *Nature Neuroscience*, *5*, 485-490.

JAIRAZBHOY, N. A. (1971). *The rāgs of North Indian music*. Middletown, CT: Wesleyan University Press.

KESSLER, E. J., HANSEN, C., & SHEPARD, R. N. (1984). Tonal schemata in the perception of music in Bali and in the West. *Music Perception*, *2*, 131-165.

KOIZUMI, F. (1977). Musical scales in Japanese music. In F. Koizumi, Y. Tokumaru, & O. Yamaguchi (Eds.), *Asian musics in an Asian perspective* (pp. 73-39). Tokyo: Heibonsha.

KRUMHANSL, C. L., & CUDDY, L. L. (2010). A theory of tonal hierarchies in music. In M. R. Jones, R. R. Fay, & A. N. Popper (Eds.), *Music perception* (pp. 51-87). New York: Springer.

KRUMHANSL, C. L., & KESSLER, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial map of musical keys. *Psychological Review*, *89*, 334-368.

MORTON, D. (1976). *The traditional music of Thailand*. Berkeley, CA: University of California Press.

PATEL, A. D. (2008). *Music, language, and the brain*. New York: Oxford University Press.

PERSICHETTI, V. (1961). *Twentieth-century harmony: Creative aspects and practice*. New York: W. W. Norton.

PRESSING, J. (1983). Cognitive isomorphisms between pitch and rhythm in world musics: West Africa, the Balkans, and Western tonality. *Studies in Music*, *17*, 38-61.

RAHN, J., & AMIOT, E. (Eds.). (2011). Tiling rhythmic canons [Special issue]. *Perspectives of New Music*, *49*(2).

TRAINOR, L. J., & TREHUB, S. E. (1993). What mediates infants' and adults' superior processing of the major over the augmented triad? *Music Perception*, *11*, 185-196.

TREHUB, S. E., SCHELLENBERG, E. G., & KAMENETSKY, S. B. (1999). Infants' and adults' perception of scale structures. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 965-975.

VOLK, A., & VAN KRANENBURG, P. (2012). Melodic similarity among folk songs: An annotation study of similarity-based categorization in music. *Musicae Scientiae*, *16*, 317-339.