A Cognitive Basis for Choosing a Solmization System [*]

Gary S. Karpinski

NOTE: The examples for the (text-only) PDF version of this item are available online at:
https://www.mtosmt.org/issues/mto.21.27.2/mto.21.27.2.karpinski.php

ABSTRACT: This article focuses on the perception and cognition involved in music listening skills as essential criteria in selecting solmization systems. Drawing on many aural key-identification studies performed by various researchers, and on the model for music perception developed by Karpinski (1990) and formalized in Karpinski (2000), it concludes that the first and most fundamental process listeners carry out while attending to the pitches of tonal music is tonic inference. In addition, a tonic is inferable without reference to a complete diatonic pitch collection. Melodies that are unambiguous with regard to their tonic might never employ all seven diatonic pitch classes, they might state those pitch classes only gradually, or they might even change the collection without changing tonic. Nonetheless, listeners are able to infer tonics quickly and dynamically under any of the above conditions. According to Butler (1992, 119), "listeners make assessments of tonal center swiftly and apparently without conscious effort" certainly well in advance of inferring or perceiving entire diatonic pitch collections. This article examines the means through which *do*-based minor movable-*do* solmization most closely models this mental process and contrasts that with *la*-based minor and its inherent inability to model the pitches of a musical passage until all seven of its diatonic members are explicitly stated (or at least implicitly present). This is not to say that *la*-based minor is ineffective, but simply that *do*-based minor most closely reflects and represents the way listeners infer tonality.

In memoriam Helen Brown

"Whereof one cannot speak, thereof one must remain silent." – Wittgenstein (1922)

[1] In North America, arguments about solmization systems burn quietly, like an underground coal fire, sporadically breaking through to the surface with copious amounts of smoke and occasional flames. These breakthroughs usually occur at scholarly conferences (sometimes even over a meal shared by otherwise amicable scholars), but hardly ever make appearances among colleagues

within any particular institution, the matter of solmization having been settled sometime much earlier, usually shortly after the department's crust had just cooled.

[2] One such argument centers on whether to use *la*-based minor or *do*-based minor movable-*do* in teaching aural skills. Simply put, *la*-based minor labels the two half steps in any diatonic collection as *mi–fa* and *ti–do* regardless of mode, whereas *do*-based minor always assigns *do* to the tonic or final regardless of collection.

[3] Larson (1993) categorized solmization systems with regard to the musical features they model. For example, *la*-based minor movable-*do* explicitly models the position of pitches within a diatonic collection (without regard to tonic), whereas *do*-based minor movable-*do* explicitly models the scale-degree functions of pitches in relation to a tonic. We can therefore refer to *la*-based minor as a *collection-oriented* solmization system and to do-based minor as a *tonic-oriented* one.[1]

[4] In an exchange that has come to embody the differences between (and the conflicts between the proponents of) *la*-based minor and *do*-based minor movable-*do* solmization, Smith (1991, 1992, 1994) and Houlahan and Tacka (1992, 1994) argued—without recourse to the scientific literature on tonal perception—over points concerning the efficacy of these systems in developing music reading and listening skills. Studies during the last century, including Multer (1978), Martin (1978), Surace (1978), Smith (1987), Lorek et al. (1991), and Taggart (1997) used similar criteria for evaluating various systems. Their results are at odds with one another and inconclusive. More recent work, such as Lorek and Pembrook (2002), and Reifinger (2012) also compared various approaches to solmization using more rigorous methodology and still found no real common ground.

[5] But if we consider the perception and cognition involved in music listening as essential criteria in selecting between these two solmization systems, we can come to some conclusions. Research has shown that when listeners attend to pitches in tonal music, the first and most fundamental process they carry out is to infer the tonic.[2] As Brown, Butler, and Jones explained it, "enculturated listeners respond to incoming context by engaging in an active discovery process oriented toward identifying an initial tonal center" (1994, 377). Butler concluded that "listeners make assessments of tonal center swiftly and apparently without conscious effort" (1992, 119). Bellman noted that this occurs "after hearing just seconds of standard tonal music" (2005, 79). Yoshino and Abe found that "at the beginning of a melody . . . listeners clearly perceive a particular key usually by the sixth to eighth tone step" (2004, 286). Butler provided a compact description of the dynamic nature of this process: "Any tone will suffice as a perceptual anchor—a tonal center—until a better candidate defeats it" (1989, 238).[3] It seems this is part of our enculturation: VanHandel and Callahan (2012) found that listeners are likely conditioned to expect the tonic pitch frequently at or near the beginnings of phrases. And Vos and Van Geenen noted that "the very first few events of a series (tones in the present case) have a privileged status in an inferential process like key finding" (1996, 187–88). Clearly, inferring the tonic is something that usually occurs rapidly, towards the beginning of most musical passages.[4]

[6] It is not just that listeners are merely speedy in inferring the tonic. They also often do so without ever hearing a complete diatonic collection. Butler observed that "listeners can and do make assessments of tonality based on small numbers of pitches" (1989, 238). As Brown, Butler, and Jones put it, "ordinary listeners are able to make tonal decisions quickly from partial and fleeting musical evidence" (1994, 377). In a series of experiments described by Matsunaga and Abe (2005, 2007, 2009, 2012), listeners inferred tonics from hexachords that might belong to two different diatonic supersets. For example, the hexachord [C–D–E–G–A–B] elicited different senses of tonic: when ordered as C5–G4–E4–A4–D4–B4, listeners (both musicians and nonmusicians) inferred a tonic of C; when ordered as B4–D4–C5–G4–E4–A4, listeners inferred a tonic of G. Cuddy (1991) used even shorter diatonic subsets—three-note groups—as stimuli in tonic-finding experiments. Her subjects rated certain pitches as exhibiting "very good" suitability as tonics, such as the pitch F for the pattern G4–C4–F4, and C for the pattern C4–E4–G4. Matsunaga and Abe found that listeners often make decisions about key after hearing only a few notes, and that they "preferred to retain a key that was interpreted earlier so that subsequent tones would be interpreted as scale tones within the retained key" (2012, 12). In other words, tonally enculturated listeners infer a tonic

early on and stick with it,[5] regardless of most of the other diatonic pitches that might or might not surround it.[6]

[7] It is important to distinguish inferring the *tonic* from inferring the *key*. The term "key finding" is often used to refer to the process of inferring both tonic and mode from a passage (e.g., "I hear this passage in A minor"). Tonic inference is the process of inferring a tonic pitch, with or without regard to its associated mode (e.g., "I hear a tonic of A in this passage").[7] In a study that used J. S. Bach's Duetto for organ (BWV 805) as a stimulus, Toiviainen and Krumhansl observed that "listeners readily found the A minor key of the key signature" after hearing only the first three pitches—A3, C4, and E3 (2003, 762). Even though listeners inferred the pitch A as tonic from these few notes, the mode is not entirely certain—it could be Dorian or Phrygian as well as minor. Vos observed that "a few tones of a tune are usually enough for the average listener to establish a scale" (2000, 403). But "a few tones" rarely imply a complete diatonic collection; instead, what listeners infer from these few tones is a tonic, with the mode to be determined if and when other members of the diatonic collection are eventually stated. Melodies that are unambiguous with regard to their tonic might state the diatonic pitch classes only gradually, they might never employ all seven diatonic pitch classes, or they might even change the collection without changing tonic. Nonetheless, we are able to infer tonics quickly and dynamically under any of these conditions. Both tonic and mode are important, but tonic inference comes first.[8]

[8] These studies and observations lead to the conclusion that, while attending to the pitches of tonal music, the first and most fundamental process listeners carry out is tonic inference. And from that we can conclude that the single most immediately knowable tonal characteristic is the tonic.[9]

[9] In fact, *all* scale degrees immediately or ultimately derive their functions from their relationships with the tonic. Without the tonic, the very notion of scale degree would not make sense. When we hear a particular pitch as, say, $\hat{7}$, we are not only perceiving the $\hat{7}$-ness of that pitch; we are also calculating (consciously or not) its position in the prevailing tonality in relation to $\hat{1}$. That "$\hat{7}$-ness" is an instance of scale-degree qualia—the sui generis qualities of the scale degrees that set them apart from one another in our experience. The qualia of various scale degrees were recognized nearly as early as the origins of tonality itself—in concepts such as the rule of the octave. Fétis discussed the character of each of the seven scale degrees, noting in particular that the "tonic appears through the absolute feeling of repose that is felt there" ([1840] 1994, 159). Huron, after informally surveying experienced Western-enculturated musicians about the "distinctive quality or character" of various scale degrees, concluded that such listeners "appear to experience broadly similar qualia" for the various scale degrees, and that nonmusicians seem to experience similar qualia (2006, 144–47). Arthur (2018) followed up on Huron's work with a formal experiment and found that her listeners were indeed fairly consistent in their ratings of scale-degree qualia. Hansberry (2017) called scale-degree qualia "a qualitatively irreducible part of tonal phenomenology" (185) and decided that "the attribution of scale-degree qualia results partially from decisions about the key of a passage" (183). As Bharucha noted, "once a tonal schema is activated by the prior musical context, subsequent tones and chords are perceived in terms of it" (1984, 490). In other words, our perception of scale degrees *depends on* our perception of the tonic.

[10] In the context of taking tonal melodic dictation (and in much tonal listening in general), inferring the tonic plays an essential and fundamental role.[10] I developed a model for the dictation-taking process that lays out the various steps listeners must carry out in order to turn sound into notation (see Karpinski 1990 and Karpinski 2000, 64–91). This model places tonic inference as the very first step listeners undertake (consciously or unconsciously) while understanding the pitches of a tonal melody. As I noted,

> All functional tonal pitch evaluations stem from a sense of the tonic. Without an ability to infer the tonic, listeners operate without the very frame of reference at the heart of tonality itself. With tonic inference, listeners can determine scale degrees, harmonic functions, modulations, and a host of other tonal features. This must be an explicit and accurately executed part of the dictation process or all subsequent pitch processing will be for naught. (2000, 82)

[11] Tonic-oriented solmization systems, such as *do*-based minor movable-*do*, most closely model this mental process because they label the tonic first and then derive the other scale degrees from there. My purpose here is to examine the means through which *do*-based minor does this, and to contrast that with *la*-based minor and its inherent inability to model the pitches of a musical passage until all seven of its diatonic members are explicitly stated (or at least implicitly present).

[12] Note that I am letting *do*-based movable-*do* represent an entire category of resting-tone-oriented systems, such as scale-degree numbers—which date back at least to Rousseau in the eighteenth century[11]—and perhaps even Mersenne's seventeenth-century extensions to hexachordal solmization.[12] Readers should feel free to read "$\hat{1}$–$\hat{2}$–$\hat{3}$" whenever I write "*do–re–mi*" in discussions of *do*-based solmization.

[13] Also note that none of this precludes choosing from various other solmization systems to develop other modes of musical thought. For instance, reading the letter names of absolute pitches (or reading in fixed-*do*, which is functionally equivalent) can develop clef-reading and transposition skills. Using *la*-based minor can develop a sensitivity to relative keys. And I am most definitely not addressing here the efficacy of various systems with regard to sight singing in general; this article addresses only listening skills. Nonetheless, if we use tonal perception as an essential criterion in choosing a solmization system, we will find that the way in which tonic-oriented solmization functions is most closely analogous to the cognitive process of tonic inference.

[14] To see (and hear) how this can play out in context, consider the opening measures of the overture to Wagner's *Flying Dutchman*, shown in **Example 1**. The collected pitch classes from all instruments through the downbeat of m. 6 constitute a mere dyad—D and A—yet listeners almost immediately infer D as tonic when hearing this passage, by virtue of D's position as the lower pitch in a perfect 5th (and upper pitch in a perfect 4th), whose members are typically perceived in a tonic-dominant relationship in the absence of other overpowering factors.[13]

[15] A tonic-oriented solmization system, such as *do*-based movable-*do* or scale-degree numbers, accommodates this mental process with ease. As soon as listeners aurally identify a tonic pitch, they can assign *do* (or $\hat{1}$) to it, and start calculating the other scale degrees from there as the music proceeds.[14] In contrast, listeners using a collection-oriented system, like *la*-based movable-*do*, must of necessity wait until they hear all seven pitch classes of a diatonic collection before they can begin to assign pitch labels without fear of being undermined. For example, while tonic-oriented listeners are already hearing *do–sol* for D and A in *The Flying Dutchman*, collection-oriented listeners might think *do–sol* if major, *la–mi* if minor, or even some other perfect-5th syllable pair in some other mode, but they cannot know for sure until they hear the entire diatonic collection.

[16] In fact, this distinction between the two systems comes into marked contrast when we address modes beyond minor and major.[15] Collection-oriented movable-*do* is often referred to as "*la*-based minor," but this is merely a convenient synecdoche. It is actually *do*-based major, *re*-based Dorian, *mi*-based Phrygian, and so on. Consider for a moment the Beatles' "Norwegian Wood," the beginning of which is given in **Example 2**.

[17] "Norwegian Wood" does not reveal all seven of its diatonic pitch classes until m. 2.2. However, listeners begin to infer a tonal center from the outset. Immediately—at m. 1.1—E asserts itself as tonic in the same way D did in the opening of *Dutchman*. Through m. 1 to the downbeat of m. 2, the "strong" first and third beats form the major triad B–G♯–E, thereby adding the mediant-tonic relationship and strengthening E's status as tonic.[16] Indeed, by the time we reach m. 2.1 there should be little doubt—among tonally acculturated listeners—that the tonic is E.

[18] But these five beats have unfolded only a hexachordal collection—E–F♯–G♯–A–B–C♯–[ ]—a subset of the diatonic collection, whose seventh diatonic member might be either D♯ or D♮. What is a solmization system to do? **Example 3** shows how a collection-oriented system (such as *la*-based minor) can model this hexachord as a subset of two different diatonic collections. First, the four-sharp diatonic collection—shown in Example 3a—places semitones (marked by angle brackets) at the points G♯–A and D♯–E. A collection-oriented system would assign the syllables *mi–fa* and *ti–do* to them respectively; the other syllables fall in place sequentially. Second, observe the differences

effected by replacing D♯ with D♮. The resulting three-sharp collection—shown in Example 3b—places semitones at the points G♯–A and C♯–D to which a collection-oriented system assigns *ti–do* and *mi–fa*, in that order. Note that every pitch is now assigned a new syllable, even though their scale-degree functions have not changed.

[19] Example 3c shows the dilemma posed by applying a collection-oriented system to the hexachordal diatonic subset presented by beats 1–5 of "Norwegian Wood." The system cannot know whether D♯ or D♮ would complete the collection. Therefore, it cannot determine the position of both semitones and as a consequence the status of all pitch-syllable relationships is in doubt. An internally valid collection-oriented system cannot say anything about these pitches until the collection is completed on the sixth beat.

[20] Now compare **Example 4**. A tonic-oriented solmization system assigns syllables on the basis of the position of the tonic. The tonic pitch E is solmized as *do* regardless of whether D♯ or D♮ completes the collection. Note the assignment of *ti* to D♯ in Example 4a and *te* to D♮ in Example 4b. [17]

[21] Example 4c shows that no dilemma exists for tonic-oriented systems when approaching the diatonic subset presented by beats 1–5 of "Norwegian Wood." Since E is clearly the tonic from the outset, we assign *do* to E, and—remembering that listeners tend to retain initial tonics as other members of a diatonic collection appear—the other syllables are assigned in relation to that tonic. The *ti/te* ambiguity is simply not a factor. That ambiguity is resolved definitively in favor of *te* on beat 6, yielding Mixolydian mode,[18] but that ambiguity need never be resolved in order to identify the tonic and thereby assign *do*-based syllables.

[22] Of course, all this happens relatively quickly in "Norwegian Wood": it is about six seconds before D♮ enters—quite long enough in cognitive terms for listeners to begin to infer the tonic, but still rather brief nonetheless. For an example that remains ambiguous for a much longer period of time, consider Guillaume de Machaut's virelai "Douce dame jolie," shown in **Example 5**. The complete diatonic collection takes much more time to unfold here. The entire refrain yields only a hexachordal diatonic subset: G–A–B♭–C–D–[ ]–F. Nonetheless, the functions of these pitches are clear from the first few moments: to speak in terms that are true to both fourteenth-century theory and twenty-first-century perception, the interplay between D and G is that of dominant and final (functionally equivalent to a tonic here), through both close proximity and ambitus.[19] However, due to the absence of letter-class E up to this point, the mode could be either Aeolian (if it were E♭) or Dorian (if it were E♮).

[23] Since one of the two diatonic semitones is not present in the refrain from "Douce dame," a collection-oriented solmization system would be unable to assign syllables unambiguously to the pitches in this passage. The hexachord G–A–B♭–C–D–[ ]–F could be solmized as *la–ti–do–re–mi–sol* or as *re–mi–fa–sol–la–do*. And neither would explicitly express the concluding function of the *finalis* in this passage.

[24] In contrast, a functional solmization system, oriented from the final or tonic, could begin to model the dynamic perceptual process listeners bring to bear on "Douce dame," starting with the opening pitches. Certainly, well before the completion of the first refrain, modern listeners will have inferred the final of G, and those using functional solmization will assign *do* to that pitch.[20]

[25] The very first pitch that follows the refrain—E♮—completes the one-flat diatonic collection: G–A–B♭–C–D–E♮–F–G. The final remains on G, but the mode is now for the first time completely unambiguous—Dorian. Listeners using collection-oriented solmization who had been treating G as *la* up to this point would need to rejigger their bearings and label it *re* to accommodate its position in the one-flat collection.

[26] Without applying any *musica ficta* principles, the rest of the verse remains in Dorian. But some might apply the *una nota super la* rule[21] and alter E to E♭ in m. 11, thereby alternating between the one-flat and the two-flat collections. In this case, an internally consistent collection-oriented system

would, by definition, be obligated to alternate along with this change, solmizing D–E♮ as *la–ti*, and D–E♭ as *mi–fa*.

[27] In comparison—since the final remains constant—a system oriented from the final could reflect this single change of pitch through a single change of syllable on the changed pitch itself: *sol–la* (D–E♮) would become *sol–le* (D–E♭). No other syllables would need to change. In a process analogous to the dynamic manner in which listeners infer the tonic and reckon scale degrees from there, a tonic-oriented solmization system labels the tonic (or final) and then labels the functions of the other members of the collection as they unfold over time.

[28] The inability of a collection-oriented solmization system to decisively model pitches until all members of a diatonic collection have appeared has direct consequences in aural skills training. **Example 6** reproduces a dictation melody taken from Benward and Kolosick's *Ear Training: A Technique for Listening* (2010). Immediately above the staff is a horizontal bar graph indicating the cumulative pitch collection at each point in time as the melody progresses. As each new pitch class is introduced, the cardinality of the collection increases — D, then D–A, then D–A–E, and so on. Below the staff is another bar graph indicating which modes that collection might belong to with a tonic or final of D.

[29] When the pitch D first enters it forms a collection of cardinality 1. In keeping with the contention in Butler 1989 that "the first tone (T1) would serve as a plausible tonal center" (239), I have (perhaps trivially) indicated that this plausible tonal center of D could potentially belong to all diatonic pitch collections from three sharps to three flats: all possible modes. When the pitch A is added at the end of m. 1, it not only strengthens D's status as tonic through the perfect 5th relationship, but also eliminates the possibility of Locrian mode since A♭ is essential to the three-flat collection. E—on the downbeat of m. 3—eliminates Phrygian, and G—on the second half of that beat—eliminates Lydian. Finally, m. 3.2 completes the entire five-member collection presented by this dictation. Note that the addition of F to the G–D–A–E tetrachord diatonically implies C♮ as well, since it is not possible for a diatonic collection to combine both F and G with any other C but C♮. Therefore, the addition of F eliminates major and Mixolydian in one stroke.[22]

[30] Although D only strengthens in its role as tonic throughout the melody, the mode to which it belongs is never completely certain. At any point from the completion of m. 1 onward, it could belong to any of six, then five, then four, then two modes. A tonic-oriented solmization system most closely models this tonic-first mode-later process as listeners track such tonal melodies: the tonic is assigned when it is perceived—early in the process—and other scale degrees are discerned only when they make themselves known, if at all. Using a tonic-oriented system, listeners would solmize the beginning of this dictation as *do–sol–sol–do*. In contrast, a collection-oriented system must wait—perhaps forever—to say anything definitive about such listening experiences. Even once all five pitch classes have been introduced, how can one know if the beginning of this little melody should be solmized *la–mi–mi–la* or *re–la–la–re*?

[31] Even if we restrict our pedagogy to the major-minor system, the problem persists. For example, were we to excerpt the first vocal passage from Mozart's *Don Giovanni* (shown in **Example 7**) to use for dictation, students would have to listen through fourteen notes until the mode is revealed as major and not minor on the fifteenth note (the A in m. 14). Leo Kraft offered up a similar passage, explicitly for dictation purposes, in *A New Approach to Ear Training* (1999), shown in **Example 8**. In this passage, the mode doesn't become clear until the *twenty-seventh* note (the C♯ in m. 5). Why should we expect students to wait so long to make a decision about what solmization syllables to apply to this passage? Indeed, why would *any* listener wait that long to start hearing tonal functions? Students using a collection-oriented solmization system would endure such a wait, but those using a tonic-oriented system would assign *do* to the tonic as soon as they infer it—regardless of mode—and immediately begin to reckon the other syllables from there.

[32] Two scale degrees—$\hat{7}$ and $\hat{6}$—merit special attention here. The presence of the leading tone can help to define the position of the *tonic*,[23] but it alone doesn't clarify the *mode*. **Example 9** excerpts eight measures from a waltz by Johann Strauss, Jr. Once again, the functions of the pitches become quite clear within moments — G and C function as dominant and tonic, and B serves as the leading

tone. Nevertheless, the mode could be either major or minor throughout these measures. Listeners cannot tell whether B ($\hat{7}$) is a member of a prevailing no-sharp/no-flat diatonic collection or a chromatically raised pitch in the three-flat collection. Compare the Strauss with the passage from Mahler's Second Symphony shown in **Example 10**. From a dictation-taker's perspective, B ($\hat{7}$) might yet again be either diatonic or chromatic since the mode is unclear (although this time the pitch is chromatic in its larger context). The sixth scale degree can have similar cognitive implications, particularly in conjunction with $\hat{7}$. **Example 11** shows the first eight measures from Brahms's song "Jungfräulein, soll ich mit euch gehn," in which listeners would quickly infer G as the tonic. Students listening to this passage while using *do*-based minor would solmize $\hat{6}$ and $\hat{7}$ (E♮ and F♯) as *la* and *ti*. But those using *la*-based minor wouldn't have enough information in these eight measures to determine if they should be heard as *la* and *ti* or *fi* and *si*. In these passages — all three suitable for melodic-dictation purposes—the mode is uncertain. And that uncertainty creates a conundrum for a collection-oriented solmization system, but is trivial for a tonic-oriented one.

[33] One final example will bring our discussion full circle while reaching back into the pedagogical materials of an earlier generation. **Example 12** shows a dictation melody taken from McHose's *Teacher's Dictation Manual* ([1948](#)). As in the opening to *Dutchman*, this melody contains nothing more than an interval-class-5 dyad—this time, A and E. Most students listening for the tonal functions of the pitches would begin to infer A as tonic before m. 1 is even completed. By the end of the dictation, A is firmly established as tonic. However, as we have seen, this dyad could potentially be a member of any diatonic mode but Locrian. This is not a quandary for *do*-based solmization, which would label the two pitch classes *do* and *sol*. However, *la*-based minor (and its associated relative modes) could potentially label them *do–sol* or *la–mi*—or even *re–la*, *mi–ti*, *fa–do*, or *sol–re*.

<p style="text-align:center">* * *</p>

[34] Having come to these conclusions, I would nevertheless like to point out that collection-oriented solmization (*la*-based minor) can be quite useful. It can be advantageous for singing collection-oriented music, like Renaissance motets (which tend to stay in one diatonic collection while visiting cadences on various pitches), or certain folk musics (such as Eastern-European and Russian melodies) that move freely from minor to relative major and back. But when it comes to modeling *functional* perception during *listening*, a tonic-oriented solmization system begins to model cognitive processes almost immediately whereas a collection-oriented system—as in Wittgenstein's aphorism—must remain silent until it becomes completely informed about the collection of an entire passage, if that collection is ever completed at all.

Gary S. Karpinski
The University of Massachusetts Amherst
Department of Music and Dance
Amherst, MA 01003
garykarp@music.umass.edu

Works Cited

Albrecht, Joshua D., and David Huron. 2014. "A Statistical Approach to Tracing the Historical Development of Major and Minor Pitch Distributions, 1400–1750." *Music Perception* 31 (3): 223–43. https://doi.org/10.1525/mp.2014.31.3.223.

Arthur, Claire. 2018. "A Perceptual Study of Scale-Degree Qualia in Context." *Music Perception* 35 (3): 295–314. https://doi.org/10.1525/mp.2018.35.3.295.

Bain, Jennifer. 2005. "Tonal Structure and the Melodic Role of Chromatic Inflections in the Music of Machaut." *Plainsong and Medieval Chant* 14 (1): 59–88. https://doi.org/10.1017/S0961137104000117.

Bellman, Héctor. 2005. "About the Determination of Key of a Musical Excerpt." In *Proceedings of Computer Music Modeling and Retrieval*, ed. Richard Kronland-Martinet, Thierry Voinier, and Sølvi

Ystad, 76–91. Springer. https://doi.org/10.1007/11751069_7.

Benward, Bruce, and J. Timothy Kolosick. 2010. *Ear Training: A Technique for Listening* (Instructor's Edition). 7th ed., rev. McGraw-Hill.

Bharucha, Jamshed J. 1984. "Anchoring Effects in Music: The Resolution of Dissonance." *Cognitive Psychology*, 16 (4): 485–518. https://doi.org/10.1016/0010-0285(84)90018-5.

Bodily, Paul M., and Dan Ventura. 2018. "Comparative Analysis of Key Inference Models for Musical Metacreation." In *Proceedings of the Sixth International Workshop on Musical Metacreation*. https://musicalmetacreation.org/mume2018/proceedings/Bodily.pdf.

Boltz, Marilyn. 1989. "Perceiving the End: Effects of Tonal Relationships on Melodic Completion." *Journal of Experimental Psychology: Human Perception and Performance* 15 (4): 749–61. https://doi.org/10.1037/0096-1523.15.4.749.

Brown, Helen, David Butler, and Mari Riess Jones. 1994. "Musical and Temporal Influences on Key Discovery." *Music Perception* 11 (4): 371–407. https://doi.org/10.2307/40285632.

Butler, David. 1989. "Describing the Perception of Tonality in Music: A Critique of the Tonal Hierarchy Theory and a Proposal for a Theory of Intervallic Rivalry." *Music Perception* 6 (3): 219–42. https://doi.org/10.2307/40285588.

—————. 1992. *The Musician's Guide to Perception and Cognition*. Schirmer Books.

Butler, David, and Helen Brown. 1994. "Describing the Mental Representation of Tonality in Music." In *Musical Perceptions*, ed. Rita Aiello, 191–212. Oxford University Press.

Chuan, Ching-Hua, and Elaine Chew. 2007. "Audio Key Finding: Considerations in System Design and Case Studies on Chopin's 24 Preludes." *EURASIP Journal on Advances in Signal Processing* 2007. https://doi.org/10.1155/2007/56561.

Cooke, Derryck. 1959. *The Language of Music*. Oxford University Press.

Cuddy, Lola. 1991. "Melodic Patterns and Tonal Structure: Converging Evidence." *Psychomusicology* 10 (2): 107–26. https://doi.org/10.1037/h0094138.

Farbood, Morwaread Mary, Gary Marcus, and David Poeppel. 2013. "Temporal Dynamics and the Identification of Musical Key." *Journal of Experimental Psychology: Human Perception and Performance* 39 (4): 911–18. https://doi.org/10.1037/a0031087.

Farbood, Morwaread Mary, Jess Rowland, Gary Marcus, Oded Ghitza, and David Poeppel. 2015. "Decoding Time for the Identification of Musical Key." *Attention, Perception, & Psychophysics* 77: 28–35. https://doi.org/10.3758/s13414-014-0806-0.

Fétis, François-Joseph. (1840) 1994. *Esquisse de l'histoire de l'harmonie: An English-language translation of the François-Joseph Fétis History of Harmony*. Translated by Mary I. Arlin. Pendragon Press.

Fujita, Tetsuya, Yuji Hagino, Hajime Kubo, and Goro Sato. 1993. *The Beatles Complete Scores*. Hal Leonard Publishing.

Gruber, Albion. 1970. "Mersenne and Evolving Tonal Theory." *Journal of Music Theory* 14 (1): 36–67. https://doi.org/10.2307/843036.

Hansberry, Benjamin. 2017. "What Are Scale-degree Qualia?" *Music Theory Spectrum* 39 (2): 182–99. https://doi.org/10.1093/mts/mtx014.

Harden, Bettie Jean. 1983. "Sharps, Flats, and Scribes: *Musica Ficta* in the Machaut Manuscripts." PhD diss., Cornell University.

Houlahan, Micheál, and Philip Tacka. 1992. "The Americanization of Solmization: A Response to the Article by Timothy A. Smith, 'A Comparison of Pedagogical Resources in Solmization Systems.'"

Journal of Music Theory Pedagogy 6: 137–51. https://jmtp.appstate.edu/reader's-response-americanization-solmization-response-article-timothy-smith-'-comparison.

—————. 1994. "Continuing the Dialogue: The Potential of Relative Solmization for the Music Theory Curriculum at the College Level." *Journal of Music Theory Pedagogy* 8: 221–25. https://jmtp.appstate.edu/readers'-comments-continuing-dialogue-potential-relative-solmization-music-theory-curriculum-college.

Huron, David. 2006. *Sweet Anticipation: Music and the Psychology of Expectation*. MIT Press. https://doi.org/10.7551/mitpress/6575.001.0001.

Huron, David, and Joshua Veltman. 2006. "A Cognitive Approach to Medieval Mode: Evidence for an Historical Antecedent to the Major/Minor System. *Empirical Musicology Review* 1 (1): 33–55. https://doi.org/10.18061/1811/24072.

Johnson, Timothy A. 1990–91. "Solmization in the English Treatises Around the Turn of the Seventeenth Century." *Theoria* 5: 42–60. https://digital.library.unt.edu/ark:/67531/metadc287868/m1/50/?q=johnson.

Kania, Dariusz, and Paulina Kania. 2019. "A Key-Finding Algorithm Based on Music Signature." *Archives of Acoustics* 44 (3): 447–57. https://doi.org/10.24425/aoa.2019.129260.

Karpinski, Gary S. 1990. "A Model for Music Perception and Its Implications in Melodic Dictation." *Journal of Music Theory Pedagogy* 4: 191–229. https://jmtp.appstate.edu/jmtp-volume-4

—————. 2000. *Aural Skills Acquisition: The Development of Music Listening, Reading, and Performing Skills in College-Level Musicians*. Oxford University Press.

—————. 2012. "Ambiguity: Another Listen." *Music Theory Online* 18 (3). https://doi.org/10.30535/mto.18.3.9.

Kraft, Leo. 1999. *A New Approach to Ear Training*. 2nd ed. W.W. Norton.

Krumhansl, Carol L. 1990a. *Cognitive Foundations of Musical Pitch*. Oxford University Press.

Krumhansl, Carol L. 1990b. "Tonal Hierarchies and Rare Intervals in Music Cognition." *Music Perception* 7 (3): 309–24. https://doi.org/10.2307/40285467.

Krumhansl, Carol L., and Lola L. Cuddy. 2010. "A Theory of Tonal Hierarchies in Music." In *Music Perception*, (*Springer Handbook of Auditory Research* 36), ed. Mari Riess Jones, Richard R. Fay, and Arthur N. Popper, 51–87. Springer. https://doi.org/10.1007/978-1-4419-6114-3_3.

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* 89 (4): 334–68. https://doi.org/10.1037/0033-295X.89.4.334.

Langhabel, Jonas, Robert Lieck, Marc Toussaint, and Martin Rohrmeier. 2017. "Feature Discovery for Sequential Prediction of Monophonic Music." In *Proceedings of the 18th International Society for Music Information Retrieval Conference*, ed. Sally Jo Cunningham, Zhiyao Duan, Xiao Hu, and Douglas Turnbull. https://dblp.org/rec/conf/ismir/2017.html.

Larson, Steve. 1993. "The Value of Cognitive Models in Evaluating Solfege Systems." *Indiana Theory Review* 14 (2): 73–116. http://www.jstor.org/stable/24045329

Leguy, Sylvette, ed. 1977. *Guillaume de Machaut: Oeuvres Complètes*, vol. 1. Le Droict Chemin de Musique.

Longuet-Higgins, H. Christopher, and Mark J. Steedman. 1971. "On Interpreting Bach." *Machine Intelligence* 6: 221–41.

Lorek, Mary Jo, and Randall G. Pembrook. 2002. "To *Doh* or not to *Doh*: The Comparative Effectiveness of Sightsinging Syllable Systems." *Journal of Music Theory Pedagogy* 14: 1–14. https://jmtp.appstate.edu/doh-or-not-doh-comparative-effectivenes-sightsinging-syllable-systems.

Lorek, Mary Jo, H. Lee Riggins, Randall Pembrook, Ken Lidge, and Laura New. 1991. "The Effect of Three Syllable Systems—Fixed *Do*, Movable *Do*, and 'Lah'—on the Sightsinging Performance of Freshmen Music Majors." Paper presented at the second annual conference of Music Theory Midwest, Kansas City, MO, May 17–19, 1991.

Ludwig, Friedrich, ed. 1926. *Guillaume de Machaut: Musikalische Werke*, vol. 1. Breitkopf & Härtel.

Martin, Louis. 1978. "Solmization: Getting the Facts Straight." *Theory and Practice* 3 (2): 21–25. https://www.jstor.org/stable/41330128.

Matsunaga, Rie, and Jun-ichi Abe. 2005. "Cues for Key Perception of a Melody: Pitch Set Alone?" *Music Perception* 23 (2): 153–64. https://doi.org/10.1525/mp.2005.23.2.153.

—————. 2007. "Incremental Process of Musical Key Identification." In *Proceedings of the 29th Annual Cognitive Science Society*, ed. Danielle S. McNamara and J. Gregory Trafton, 1277–82. Cognitive Science Society. https://escholarship.org/uc/item/9nr5c65k

—————. 2009. "Do Local Properties Function as Cues for Musical Key Perception?" *Japanese Psychological Research* 51 (2): 85–95. https://doi.org/10.1111/j.1468-5884.2009.00391.x.

—————. 2012. "Dynamic Cues in Key Perception." *International Journal of Psychological Studies* 4 (1): 3–21. https://doi.org/10.5539/ijps.v4n1p3.

McHose, Allen Irvine. 1948. *Teachers Dictation Manual*. Appleton-Century-Crofts.

Multer, Walt. 1978. "Solmization and Musical Perception." *Theory and Practice* 3 (1): 29–51. https://www.jstor.org/stable/41330411.

Oxford English Dictionary Online. 2020. s.v. "Tonality, n." Oxford University Press. www.oed.com/view/Entry/203142. Accessed 30 August 2020.

Prince, Jon B., and Mark A. Schmuckler. 2014. "The Tonal-Metric Hierarchy: A Corpus Analysis." *Music Perception* 31 (3): 254–70. https://doi.org/10.1525/mp.2014.31.3.254.

Rameau, Jean-Philippe. 1737. "Génération Harmonique, ou Traité de Musique Théorique et Pratique." In *The Complete Theoretical Writings of Jean-Philippe Rameau*, vol. 3, ed. Erwin R. Jacobi. American Institute of Musicology.

Reifinger, James L., Jr. 2012. "The Acquisition of Sight-Singing Skills in Second-Grade General Music: Effects of Using Solfège and of Relating Tonal Patterns to Songs." *Journal of Research in Music Education* 60 (1): 26–42. https://doi.org/10.1177/0022429411435683.

Schoenberg, Arnold. (1934) 2010. "Problems of Harmony." In *Style and Idea*, ed. Leonard Stein, trans. Leo Black, 268–87. University of California Press.

Schrade, Leo, ed. 1956. *The Works of Guillaume de Machaut: Second Part*. Vol. 3 of *Polyphonic Music of the Fourteenth Century*. Editions de l'Oiseau-Lyre.

Shanahan, Daniel. 2017. "Musical Structure: Tonality, Melody, Harmonicity, and Counterpoint." In *The Routledge Companion to Music Cognition*, ed. Richard Ashley and Renee Timmers, 141–51. Taylor & Francis. https://doi.org/10.4324/9781315194738-12.

Smith, Timothy A. 1987. "Solmization: A Tonic for Healthy Musicianship." *The Choral Journal* 28 (1): 16–23. https://www.jstor.org/stable/23547707.

—————. 1991. "A Comparison of Pedagogical Resources in Solmization Systems." *Journal of Music Theory Pedagogy* 5: 1–23. https://jmtp.appstate.edu/comparison-pedagogical-resources-solmization-

[systems](#).

——————. 1992. "Liberation of Solmization: Searching for Common Ground." *Journal of Music Theory Pedagogy* 6: 153–68. https://jmtp.appstate.edu/author's-reply-liberation-solmization-searching-common-ground.

——————. 1994. "Ending the Dialogue: Imaginary Solutions are No Solution." *Journal of Music Theory Pedagogy* 8: 227–30. https://jmtp.appstate.edu/readers'-comments-ending-dialogue-imaginary-solutions-are-no-solution.

Surace, Joseph A. 1978. "'Transposable Do' for Teaching Aural Recognition of Diatonic Intervals." *Theory and Practice* 3 (2): 25–27. https://www.jstor.org/stable/41330129.

Taggart, Bruce. 1997. "Sight Singing Schubert: A Study in Solfège." *Journal of Music Theory Pedagogy* 11: 75–98. https://jmtp.appstate.edu/sight-singing-schubert-study-solfege.

Temperley, David. 1999. "What's Key for Key? The Krumhansl-Schmuckler Key-Finding Algorithm Reconsidered." *Music Perception* 17 (1): 65–100. https://doi.org/10.2307/40285812.

——————. 2007a. *Music and Probability*. MIT Press. https://doi.org/10.7551/mitpress/4807.001.0001.

——————. 2007b. "The Tonal Properties of Pitch-Class Sets: Tonal Implication, Tonal Ambiguity, and Tonalness." *Computing in Musicology* 15: 24–38.

Temperley, David, and Trevor de Clercq. 2013. "Statistical Analysis of Harmony and Melody in Rock Music." *Journal of New Music Research* 42 (3): 187–204. https://doi.org/10.1080/09298215.2013.788039.

Toiviainen, Petri, and Carol L. Krumhansl. 2003. "Measuring and Modeling Real-Time Responses to Music: The Dynamics of Tonality Induction." *Perception* 32 (6): 741–66. https://doi.org/10.1068/p3312.

VanHandel, Leigh, and Michael Callahan. 2012. "The Role of Phrase Location in Key Identification by Pitch Class Distribution." In *Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music*, ed. Emilios Cambouropoulos, Costas Tsougkras, Panayiotis Mavromatis, and Konstantinos Pastiadis, 1069–73. http://icmpc-escom2012.web.auth.gr/files/papers/1069_Proc.pdf

Vos, Piet G. 1999. "Key Implications of Ascending Fourth and Descending Fifth Openings." *Psychology of Music* 27 (1): 4–17. https://doi.org/10.1177/0305735699271002.

——————. 2000. "Tonality Induction: Theoretical Problems and Dilemmas." *Music Perception* 17 (4): 403–16. https://doi.org/10.2307/40285826.

Vos, Piet G., and Erwin W. Van Geenen. 1996. "A Parallel-Processing Key-Finding Model." *Music Perception* 14 (2): 185–223. https://doi.org/10.2307/40285717.

White, Christopher W. 2018. "Feedback and Feedforward Models of Musical Key." *Music Theory Online* 24 (2). https://doi.org/10.30535/mto.24.2.4.

Wittgenstein, Ludwig. 1922. *Tractatus Logico-Philosophicus*. Harcourt, Brace & Company.

Yoshino, Iwao, and Jun-ichi Abe. 2004. "Cognitive Modeling of Key Interpretation in Melody Perception." *Japanese Psychological Research* 46 (4): 283–97. https://doi.org/10.1111/j.1468-5584.2004.00261.x.

Footnotes

1. *La*-based minor can also be labeled a *relative* approach and *do*-based minor a *parallel* one (see Karpinski 2000, 86).
Return to text

2. This does not mean that it is clear exactly *how* listeners do this. As Bodily and Ventura remarked, "there is a profound irony in the contrast between knowing how and being able to explain how to infer the key of a musical selection. Expert musicians routinely and accurately perform this task. However the description for their methodology is often inexact" (2018, 2).
Return to text

3. Brown, Butler, and Jones dubbed this the "primacy hypothesis" (1994, 372).
Return to text

4. In using the term "rapidly" here, I am referring to the rate at which listeners infer tonics at the beginnings of typical musical passages. For research on how rapidly pitch sequences can be played while listeners can still infer tonics, see Farbood, Marcus, and Poeppel 2013 and Farbood et al. 2015.
Return to text

5. Matsunaga and Abe labeled this propensity "perceptual inertia" (2012, 3).
Return to text

6. Similarly, most of the so-called "key-finding" algorithms developed in recent decades, which are often designed to model how listeners infer both tonic and mode (i.e., "find the key"), operate on small handfuls of pitches. In her groundbreaking series of experiments, Carol Krumhansl tested her tonal-hierarchy algorithms on the first four notes of various works (see, for example, Krumhansl 1990a, 81–96). Others have followed suit. Kania and Kania (2019) applied their model to the first four notes of each of the preludes and fugues in Bach's *Well-Tempered Clavier*. Temperley (1999) divided passages into many small segments and had his model judge the key of each segment in isolation; the model began to make judgments on openings with as few as three pitch classes. The key-finding algorithm described in Chuan and Chew 2007 determined the tonic with rather high accuracy within the first few seconds of processing an audio signal. For reviews of various approaches to automated key finding, see Temperley 2007a (49–56) and Shanahan 2017 (143–45). For a summary of Krumhansl's work on tonal hierarchy, see Krumhansl and Cuddy 2010. Readers should keep in mind that, although such algorithms are designed to replicate the results of human tonic inference, it's possible they don't model the actual processes we use to infer tonics.
Return to text

7. Some writers use slightly different terms for the same phenomena. For example, Butler and Brown refer to *key* or *key level* as the "central and most important pitch" and the *mode* as "the more specialized set of relationships within that key level" (1994, 197).
Return to text

8. Once again, computer modeling offers some interesting parallels: Chuan and Chew (2007) found that their key-finding algorithm frequently made parallel-key errors during the first few seconds of an input signal, leading to a condition in which the tonic is determined but the mode is not. The algorithm used by Langhabel et al. includes an approach designed "to estimate the tonic based on all tones heard so far . . . but still ignore the mode" (2017, 651). While investigating rock songs, Temperley and de Clercq made the following observation: "We found in creating our corpus that it was often quite problematic to label songs as major or minor. . . . Thus, we simply treat a 'key' in rock as a single pitch-class" (2013, 194). In other words, in the language of tonic-oriented solmization, they focused on *do* regardless of whether $\hat{3}$ is *mi* or *me*. Inferring the tonic of a passage, then, is more straightforward and immediate than inferring its mode. Compare this to certain key-finding algorithms (e.g. Longuet-Higgins and Steedman 1971) that infer key *signatures* (i.e., complete diatonic collections) by waiting for a complete diatonic collection before making an assessment. Bodily and Ventura concluded that such systems "inherently fail to provide a way of distinguishing between a major key and its relative minor" (2018, 4).
Return to text

9. Indeed, the essence of tonality itself is bound up in the referentiality of the tonic pitch. The *Oxford English Dictionary Online* (2020) describes the etymology of the word *tonality* as the abstract-noun version of *tonal*, derived from the Latin *tonus*, which refers to a single pitch—i.e., the tonic. Huron offered one definition of *tonality* as "a system for interpreting pitches or chords through their relationship to a reference pitch, dubbed the *tonic*" (2006, 143). Cooke, in discussing scale degrees in the C-major scale, remarked that "there is, of course, a tension pulling every note back to the fundamental [tonic], C" (1959, 47). Schoenberg described tonality as "the particular way in which all tones relate to a fundamental tone [the tonic]" ([1934] 2010, 270.). Rameau characterized the tonic as "the center of the mode, to which all our wishes tend" (1737, 108–9).
Return to text

10. This assumes that dictation is presented in a way that does not provide or reveal the tonic and mode prior to the sound of the dictation itself. For more on this, see Karpinski 2000 (92–98).
Return to text

11. For a brief summary of solmization systems that model scale degrees, see Martin 1978 (23).
Return to text

12. See Gruber 1970. See also Johnson 1990–91 for a discussion of even earlier attempts to expand or replace hexachordal solmization to account for complete diatonic collections.
Return to text

13. Vos referred to the ascending 4th/descending 5th as "powerful intervallic information for the key" (2000, 414). In an earlier study, he concluded that "if a Western tonal composition opens melodically with an ascending fourth or a descending fifth ('4/5 opening'), then the second tone is the tonic of the composition's key" (Vos 1999, 4). He also surveyed a handful of eighteenth- and nineteenth-century treatises that cite or allude to this effect. Also note Longuet's and Higgins's (1971) "tonic-dominant preference rule," which privileges scale degrees $\hat{1}$ and $\hat{5}$ at the outset of a passage.
Return to text

14. In this sense, inferring the tonic is mostly a "feedforward" process: once a tonic pitch is inferred, all subsequent pitches are heard as scale degrees in relation to that tonic. Feedback occurs only if a listener has initially misidentified the tonic or the music begins to modulate. For more on feedback and feedforward models of key finding, see White 2018.
Return to text

15. See Huron and Veltman 2006 for an investigation of "mode profiles" similar to the key profiles developed by Krumhansl and Kessler (1982) and others. Huron and Veltman's work uses a plainchant corpus selected randomly from the *Liber usualis*. For work on mode profiles in later music, see Albrecht and Huron 2014.
Return to text

16. In the absence of other stronger factors, pitches that form a triad tend to lead listeners to conclude that the triad's root is the tonic. Matsunaga and Abe (2012) presented two-, three-, and four-pitch stimuli to subjects, and observed that the members of the tonic triad "exert greater influence on participants' interpretations [of tonic] than do other scale tones" (8). See also Cuddy 1991. With regard to the metric placement of these structural key-defining pitches, see Prince and Schmuckler 2014, which demonstrated that "tonally stable pitch classes were more likely to occur on metrically stable temporal positions" in a corpus of common-practice-period compositions (263).
Return to text

17. This approach has its origins as least as far back as Mersenne's addition of the syllables "*ci*" and "*bi*" for the whole and half steps above *la*, respectively. (Once again, see Gruber 1970, for more on this.)
Return to text

18. See Temperley and de Clerq 2013 for a discussion of the use of Mixolydian (here in the verse) and Dorian (in the bridge) in the context of the "pentatonic union scale" in "Norwegian Wood" (202–3). Note that the tonic remains E regardless of whether the mode is Mixolydian or Dorian.
Return to text

19. In addition, if one were to alter F to F♯ through *musica ficta* at the end of m. 7, this would further strengthen G as a final. Bain observed that even if the principles of *ficta* would not require an F♯, G's central status is well established: "Musical elements in the song (other than F♯) set up *G* very clearly as a tonal centre. In *Douce dame jolie* the opening leap of a descending fifth directs the ear *immediately* [emphasis added], setting up two focal pitches that permeate the song, *d* and *G*" (2005, 78).
Return to text

20. Of course, this differs from hexachordal solmization, which is not a functional system and is no longer in use as a pedagogical tool.
Return to text

21. As Schrade (1956) does, but not Ludwig (1926) nor Leguy (1977). See also Harden, who states that the rule "is now generally accepted to be absent from medieval sources. . . . [and] it certainly does not apply to his [Machaut's] compositions" (1983, 53).
Return to text

22. For more on the cognitive implications of such diatonic subsets, see Karpinski 2012 (4.1–5.2), Temperley 2007b (34–37), and Butler and Brown 1994, pp. 200–201, in particular.
Return to text

23. The leading-tone to tonic relationship is recognized as a functional part of key induction. Experiments in Boltz 1989 found that $\hat{7}$–$\hat{1}$ resolutions elicited the strongest feelings of "completeness." Krumhansl noted the great potential of the minor second as "a cue in key-finding" (1990b, 320). Brown, Butler, and Jones observed that "ascending semitone motion . . . is interpretable as a small-scale reference to leading tone to tonic" (1994, 405).
Return to text

Copyright Statement