A JOURNAL OF THE SOCIETY FOR MUSIC THEORY

Rethinking Aural Skills Instruction through Cognitive Research: A Response

Elizabeth West Marvin

REFERENCE: https://mtosmt.org/issues/mto.21.27.2/mto.21.27.2.karpinski.html

REFERENCE: https://mtosmt.org/issues/mto.21.27.2/mto.21.27.2.chenette.html

REFERENCE: https://mtosmt.org/issues/mto.21.27.2/mto.21.27.2.gates.html

KEYWORDS: aural skills, pedagogy, curriculum, music cognition, solfège, key finding, auditory imagery

ABSTRACT: This essay responds to three papers appearing in this issue that relate music-cognitive research to aural skills pedagogy. Gary S. Karpinski focuses on tonic inference as support for *do*-based minor solfège pedagogy. My discussion supports this position, with evidence from key-profile experiments and corpus analyses. Timothy Chenette proposes a perceptually based learning sequence for aural skills instruction. He sketches a model curriculum, to which I propose a staffing solution and offer a research-based challenge: the high-voice superiority principle. Finally, Sarah Gates considers what the cognitive sciences can tell us about auditory imagery. I offer classroom strategies that take advantage of motor-area activation in the brain.

Received September 2020

Volume 27, Number 2, June 2021

[1.1] The three preceding essays address an important question often posed to music theorists whose research takes a cognitive bent: how can the explosive growth of research in music cognition inform aural skills pedagogy in a practical way? These authors attempt to answer that question. Gary S. Karpinski focuses on a well researched skill that is fundamental to musical transcription: tonic inference from an aural stimulus. Timothy Chenette invites us to consider a perceptually based learning hierarchy to shape aural skills curricula, rather than a topics-based sequence that mirrors the organization of harmony courses. Sarah Gates considers what the cognitive sciences can tell us about another skill fundamental to aural musicianship: auditory imagery, or the inward hearing of music.

[1.2] We begin with Karpinski's paper, which revisits the well-trod question of *do*-based vs. *la*-based minor in movable-*do* solfège pedagogy as viewed through the lens of music-cognitive research. He argues first that tonic inference is a fundamental process of music cognition. Second, he demonstrates via musical examples that the full diatonic collection is not required for tonic to be perceived. Third, he shows how the pitch collection of a melody may unfold through time, gradually revealing its mode or scale, and he highlights the attendant problems this raises for solmization in the *la*-based system. As he points out, listeners don't know whether "Norwegian Wood" is major or Mixolydian until nearly the end of its opening phrase; yet we can infer tonic earlier on, and therefore know that the tune begins on *sol* (2021, [17]–[22]). Once we know where *do* is, we can employ inflections within *do*-based minor to assign solmization to the tune, no matter the scale or mode. This is not so with *la*-based minor, he argues, which requires that the full collection be heard to assign the correct mode and solfège syllable for its final (and *re* for Dorian, *mi* for Phrygian, etc.). If instructors use solmization as a tool for melodic dictation—using, for example, the "proto-notation" method advocated by Karpinski 2000—then finding *do* and teaching students to assign solfège quickly and accurately is paramount.

[1.3] But what does music cognition tell us about extracting the tonic? How do we come to perceive this scale degree as most stable, as a perceptual anchor? Karpinski's discussion begins with the cognitive principles proposed by Brown, Butler, and Jones 1994 and Butler 1989, in which the authors—building on Browne 1981—argue that the rare intervals of the diatonic collection (particularly the tritone and semitone) assist with aural position-finding within the scale. I have no doubt that this is one method by which tonic inference occurs. Others are at play as well, as Karpinski demonstrates in melodies without a complete diatonic collection and indeed without tritones and semitones, as in the opening to Wagner's *The Flying Dutchman* (2021, Example 1). What he doesn't mention is that Butler (1990) once engaged in a heated published exchange with music psychologist Carol Krumhansl over the question of how listeners determine the tonic when listening to music (Krumhansl 1990). Krumhansl's critique relies on data from a series of experiments that use the probe-tone technique to measure listener perceptions of the relatedness of each pitch class to a given tonal context, by rating them from 1 to 7 (Krumhansl and Shepard 1979; Krumhansl and Kessler 1982). Results are shown as key profiles, which outline a tonal hierarchy—with the scale degrees of the tonic triad rated highest by listeners, the remaining diatonic tones next highest, and non-diatonic tones lowest.

[1.4] Importantly, Krumhansl (2001) demonstrates, through corpus analysis, a strong correlation between listener ratings of each scale degree and its frequency in Western tonal music repertoire. From this, we can extrapolate that listeners' scale-degree perceptions are implicitly learned, from an unconscious tally of how often each scale degree occurs in the music to which we are exposed in daily life.[(1)] High listener ratings for $\hat{1}$ and $\hat{5}$ are arguably related to the higher frequency with which $\hat{1}$ and $\hat{5}$ are found in the corpus. Further, the profiles for parallel major and minor keys are strikingly similar, with $\hat{1}$ and $\hat{5}$ receiving the highest ratings, and small differences appearing only for modal scale degrees (for example, higher ratings for $\flat\hat{3}$ than $\hat{3}$ in minor keys). I think these key profiles bolster the case for *do*-based minor. First, they display ratings gathered from both musicians and nonmusicians, showing how widespread the perception is that there is a single, focal scale degree that can be identified aurally in parallel major and minor keys. This implicit knowledge is best modeled by a system that gives $\hat{1}$ the label do in both modalities. Implicit learning also tracks transitional probabilities *between* scale degrees—for example, how often $\hat{7}$ goes to $\hat{1}$. This leads, as Huron 2006 beautifully shows, to perceptions of scale-degree *qualia*—such as the "leading" quality of leading tones—which surely should be *ti*-*do* in both major and minor. While Karpinski discusses scale-degree *qualia* with data from verbal descriptions and listener ratings, he does not include the corroborating evidence from tracking transitional probabilities between scale degrees in musical corpora, which strengthen his case.

[1.5] Implicit in this pedagogical discussion is the assumption that music principles studied here derive from Western art music. Even "Norwegian Wood" demonstrates structural principles not too far removed from the Western art-music canon that forms the basis of many music curricula. In an era when teachers are actively seeking to diversify the music considered in the theory classroom, we may need to consider how the principles we teach are dependent upon particular repertoires, and how pedagogical and cognitive strategies might differ if that repertoire were changed. For example, key-finding strategies might differ if popular music were included more systematically in our curricula. As corpus data from Temperley and de Clercq 2013 suggest, aural determination of major or minor mode might not be so clear in popular repertories, as Karpinski notes (2021, [6]). In recent popular music based on repeated harmonic loops, tonic inference can become quite challenging when loops suggest more than one possible tonic chord and thus an ambiguous key (see, for example, Richards 2017 and Clendinning 2018).[2] What other adjustments to our aural skills pedagogy might be necessitated by an infusion of popular and non-Western musics into the undergraduate curriculum?

[2.1] Timothy Chenette invites us to ponder which skills are "truly aural." He suggests that truly aural skills are those that directly engage auditory working memory and attentional control.[3] Auditory working memory is the mental faculty that allows listeners to hold a musical idea in memory and to manipulate it—for example, by singing back with solfège syllables, altering a melody's mode or rhythm in the mind, or remembering the beginning of a melody to transcribe it (while blocking out its continuation). Auditory attentional control comes into play when musicians direct their attention selectively to soprano or bass for harmonic dictation, or isolate the violin melody from an orchestral texture, or attend to cadences to identify phrase types. A number of researchers have compared musicians and nonmusicians in auditory tasks to assess their working memory and attentional control. As an example, Schulze, Mueller, and Koelsch (2011) asked participants to listen to five-note pitch patterns, half of which featured a prominent tonic triad and half of which were more ambiguous as to key. Listeners were asked to rehearse the melodies mentally for four to six seconds, and then were given a probe tone to identify as belonging to the sequence or not belonging. During the task, participants were scanned with fMRI neuroimaging. Musicians outperformed nonmusicians in both conditions; musicians (but not nonmusicians) performed significantly better on the tonal stimuli than the tonally ambiguous patterns. In addition —for musicians only—mental rehearsal of the patterns showed functional differences in brain areas known for strategy-based processing—again, in the tonal patterns, but not the ambiguous ones. In other words, during mental rehearsal, the musicians were actively seeking a strategy for encoding and remembering. Such working-memory strategies may well have been tonic-inference, scale-degree identification, and so on—strategies they were less able to apply to the tonally ambiguous stimuli. I find that this cognitive search for structure argues in favor of what Chenette (perhaps polemically) calls a "logical" ordering in aural skills curricula, at least during the early stages of training, when students are learning to hear and name fundamental patterns of tonal and rhythmic structure. But his point is well taken that we tend to ignore other auditory working-memory and attentional-control tasks that do not fit so neatly into a theory curriculum, such as listening for blend in an ensemble, memorizing canonical literature, identifying timbres and spatial characteristics of sound, and so on.

[2.2] Indeed, I am very much struck by the important musical skills Chenette has captured in his cross-disciplinary diagram of aural skills course topics—skills applicable to historical performance practice, jazz, ethnomusicology, music therapy, audio engineering, and more (2021, Example 8). I would suggest adding a popular-music area to the diagram, especially as curricula are diversifying repertoire and schools are creating commercial and popular-music majors. While some of the skills would be similar to those for jazz (memorization of canonical literature and style-specific improvisation) and for audio engineering (listening for sound quality, identifying filters, timbre analysis), the literature, progressions, and styles would differ. Everything on this chart is an aural skill truly relevant to the careers for which our students are preparing. Yet many of these skills are missing from our aural skills classes, and with musical relevance comes student buy-in and intrinsic motivation to learn (Ryan and Deci 2000; Marvin 2018). So, I am very sympathetic to the point he is making. I like the idea of incorporating logical curriculum units like "developing internal hearing" or "tuning and intonation." But I would like to ask a polemical question myself: why is it that the music theory department owns aural skills training? Perhaps the final semester of an aural skills curriculum is the time to consider offering discipline-specific aural skills classes, designed and taught by faculty in those disciplines. Why not have the conducting faculty offer a skills class that includes multi-part error detection, clef transposition, polyphonic listening for blend and balance, and memorization and transcription of canonical literature? Why not have the ethnomusicology and jazz departments offer a skills class focused on the types of transcription and genre/style identification necessary for those disciplines? Why not have keyboard faculty teach sing-and-play sight reading, style-specific improvisation, error detection while accompanying a soloist, and instruction on hearing and conveying form? One can envision many more such discipline-specific aural skills courses, and by the fourth or fifth semester of study, students will be ready for them.

[2.3] Chenette's Examples 2–5 accompany an extended discussion of harmonic dictation. He points out that this task "involves multiple simultaneous 'voices,' and it is well established that human attention can have a high degree of focus on only one object at a time" (2021, [2.4]). He goes on to focus on bass-line dictation as the preferred focus of attention. As widespread as this strategy is in aural skills classrooms, it flies in the face of music-cognitive research, which has reliably shown a "high-voice superiority" in homophonic textures (Marie et al. 2012; Marie and Trainor 2014; and Trainor et al. 2014). Researchers have shown through EEG studies that two-voice musical textures form two separate memory traces in the brain, and that the higher of the voices shows a larger response to changed pitches when compared with the lower voice—even in children as young as seven months old, suggesting that the high-voice superiority effect may be innate. Only years of practice on a bass-range musical instrument can mitigate this effect somewhat (Marie et al. 2012). Trainor et al. 2014 hypothesize that the high-voice superiority effect may be a by-product of the physiology of the auditory system. Huron 2016 suggests that this effect may be a feature of auditory masking. What does this mean for harmonic dictation? While dictation of the lowest voice does indeed provide necessary information for a Roman numeral analysis, students may struggle to focus auditory attention on the bass line, as this works against cognitive principles, acoustics, and possibly physiology. Further, upper voices tend to feature simpler stepwise motion while bass lines feature more skips and leaps. Together, these findings suggest that Chenette's perceptually based learning hierarchy might ask students first to focus on the soprano voice in homophonic textures, before turning their attention to the bass and Roman numerals or functional symbols.

[3.1] We turn finally to auditory imagery, the ability to imagine sound that is not actually present. Sarah Gates asks whether aural skills classes—for all their talk of "inner hearing" and "audiation"—actually improve students' ability to imagine sound.(4) To answer this question she introduces the Bucknell Auditory Imagery Scale (or BAIS) developed by psychologist Andrea Halpern (2015) and discusses expertise acquisition using a long-term working memory model. BAIS is a psychometric test that presents auditory scenes to be imagined and manipulated, with participants assessing their success at each task. Gates describes her ongoing research using BAIS as a pre- and post-semester measure for aural skills students and discusses ways teachers might focus explicitly on developing students' auditory imagery. One way, which she calls "content acquisition," is to teach students typical musical patterns and melodic-harmonic schemas to recognize and label—patterns that they will learn to hear inwardly. Another is semantic encoding, assignment of solfège or scale-degree numbers to what is heard, to improve the quality of the imagined sound. A third involves cuing strategies for auditory imagery, such as subvocalization (covert tension in the larynx that simulates singing, but in silence) and multimodal associations (such as imagined instrument playing, silent whistling, or imagined notation).

[3.2] These multimodal associations are neurologically based: neuroimaging shows a strong link between motor planning and auditory imagery. For example, Bangert et al. (2006) and Meister et al. (2004) demonstrate that a pianist imagining a piano melody will engage the motor areas of the brain necessary for its performance. It follows that moving the fingers may assist in creating a more robust auditory image, and Gates suggests that incorporating keyboard study into the aural skills classroom will strengthen students' multimodal sound-to-action connections as well as their auditory imagery (2021, [2.6]). Similarly, Lévêque and Schön (2015) showed greater activations in motor areas responsible for singing when listeners heard sung melodies versus melodies in other timbres. This finding resonates well with the classroom challenge of students who have difficulty singing back pitch patterns or melodies played on the piano. For some students, performance improves when those melodic patterns are sung to them (rather than played). It may be simpler for them to sing back patterns from a sung model because the brain's motor areas for singing have been activated while listening.

[3.3] Pfordresher, Halpern, and Greenspon (2015) have developed a multi-modal model that represents automatic associations between imagery and action. For singing, this mapping is between pitch height and regulation of vocal-fold tension. They posit that poor singers have deficits in the mapping between imagery and motor action; in particular, these singers tend to drift toward a preferred vocal range and a "comfort pitch." As a practical application for the classroom, aural skills teachers should consider spending class time improving vocal technique: vocalizing our classes as one might warm up a choir, encouraging good posture and breath control, and working to expand non-singers' vocal ranges. These authors also note the "somewhat puzzling" finding that poor singers have more difficulty singing back a single repeated pitch than they do a short pitch pattern (2015, 247). This should come as no surprise to an aural skills teacher, however. Isolated pitches have no musical context and thus cannot be assigned labels (like numbers or solfège) to assist with students' auditory imagery. Three- and four-note patterns that arpeggiate tonic, dominant, and subdominant functions are the contextual building blocks of Edwin Gordon's music learning sequence (2012). Importantly, each pattern is learned with the associated solfège syllables, in what Gates calls "semantic encoding of musical content" to improve auditory imagery and to free up memory to assist with multitasking, as would be necessary in improvisation (2021, [2.4] and [2.8]). In the classroom, students who are non-singers may have difficulty mapping musical patterns from non-vocal timbres (e.g., piano or experimental sine tones) to their voices, as suggested above by Lévêque and Schön 2015. Sung short pitch patterns, on the other hand, give students a tonal context that facilitates the transfer from imagery to motor action. Taken together, these studies suggest that motor activity—from subvocalization to instrumental fingering—should not be discouraged in the classroom as a "cheat" that substitutes for true auditory imagery; rather, this strategy should be embraced as a mimetic way of knowing, to enhance auditory imagery.

[3.4] A quarter century has passed since Butler and Lochstampfor considered the "bridges unbuilt" between music cognition and aural skills pedagogy (1993, 1). As the essays published here show, much progress has been made on both sides of the fence—both by music-cognitive researchers designing experiments with ecological validity and relevance to real music making, and by music theory pedagogues creating music curricula that respond to findings in the cognitive sciences. Looking back at the 1993 essay, I see how prescient these authors were as to the questions we are addressing today. They begin with two pedagogical vignettes that touch upon solfège methodologies, including *la*-based minor, as does Karpinski in the present issue. They also address auditory imagery—as Gates does here—asking "is there a motor-learning link, as well as a memory link at the cognitive level between the pedagogical devices of numbers or syllables and the clarity and strength of our mental images of tonal pitch relationships?" (Butler and Lochstampfor 1993, 8). Further, as the basis for planning aural skills curricula from the ground up, they pose questions strikingly similar to Chenette, such as: "Are solfege and dictation going to continue to be the staples in our students' aural training regimen? . . . What could augment or replace them? . . . What about the other rationales for solfege drill, such as sight-reading practice, exposure to the performance literature, development of musical phrasing and articulation habits and the like?" (16). While many connections between music cognition and aural skills pedagogy remain to be forged, I am encouraged by the contributions published here and by the growth of music-cognitive research within music theory.(5) I look forward to continued interactions and collaborations between our disciplines, to our mutual benefit.

Elizabeth West Marvin
Eastman School of Music of the University of Rochester
26 Gibbs Street
Rochester, NY 14604
bmarvin@esm.rochester.edu

Works Cited

Bangert, Marc, Thomas Peschel, Gottfried Schlaug, Michael Rotte, Dieter Drescher, Hermann Hinrichs, Hans-Jochen Heinze, and Eckart Altenmüller. 2006. "Shared Networks for Auditory and Motor Processing in Professional Pianists: Evidence from fMRI Conjunction." *Neuroimage* 30 (3): 917–26. https://doi.org/10.1016/j.neuroimage.2005.10.044.

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.

Browne, Richmond. 1981. "Tonal Implications of the Diatonic Set." *In Theory Only* 5: 3–21.

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.

——————. 1990. "Response to Carol Krumhansl." *Music Perception* 7 (3): 325–38. https://doi.org/10.2307/40285468.

Butler, David, and Mark Lochstampfor. 1993. "Bridges Unbuilt: Comparing the Literatures of Music Cognition and Aural Training." *Indiana Theory Review* 14 (2): 1–17. https://www.jstor.org/stable/24045326.

Castellano, Mary A., Jamshed J. Bharucha, and Carol L. Krumhansl. 1984. "Tonal Hierarchies in the Music of North India." *Journal of Experimental Psychology: General* 113 (3): 394–412. https://doi.org/10.1037/0096-3445.113.3.394.

Chenette, Timothy. 2018. "Reframing Aural Skills Instruction based on Research in Working Memory." *Journal of Music Theory Pedagogy* 32: 3–20.

——————. 2021. "What Are the Truly Aural Skills?" *Music Theory Online* 27 (2). https://mtosmt.org/issues/mto.21.27.2/mto.21.27.2.chenette.html.

Clendinning, Jane Piper. 2018. "Despacito (2017) through a Music Analyst's Lens: Close Analysis of a Global Hit." Paper presented at CITYMAC (Society for Music Analysis Conference), London, UK, July 6, 2018.

Gates, Sarah. 2021. "Developing Musical Imagery: Contributions from Pedagogy and Cognitive Science." *Music Theory Online* 27 (2). https://mtosmt.org/issues/mto.21.27.2/mto.21.27.2.gates.html.

George, Elyse M., and Donna Coch. 2011. "Music Training and Working Memory: An ERP Study." *Neuropsychologia* 49 (5): 1083–94. https://doi.org/10.1016/j.neuropsychologia.2011.02.001.

Gordon, Edwin E. 2012. *Learning Sequences in Music: Skill, Content, and Patterns*. GIA Publications.

Halpern, Andrea R. 2015. "Differences in Auditory Imagery Self-Report Predict Neural and Behavioral Outcomes." *Psychomusicology: Music, Mind, and Brain* 25 (1): 37–47. http://dx.doi.org/10.1037/pmu0000081.

Hansen, Mads, Mikkel Wallentin, and Peter Vuust. 2013. "Working Memory and Musical Competence of Musicians and Non-Musicians." *Psychology of Music* 41 (6): 779–93. https://doi.org/10.1177/0305735612452186.

Hubbard, Timothy L. 2010. "Auditory Imagery: Empirical Findings." *Psychological Bulletin* 136 (2): 302–29. https://doi.org/10.1037/a0018436.

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

——————. 2016. *Voice Leading: The Science Behind a Musical Art*. MIT Press. https://doi.org/10.7551/mitpress/9780262034852.001.0001.

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

——————. 2021. "A Cognitive Basis for Choosing a Solmization System." *Music Theory Online* 27 (2). https://mtosmt.org/issues/mto.21.27.2/mto.21.27.2.karpinski.html.

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

——————. 2001. *Cognitive Foundations of Musical Pitch* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195148367.001.0001.

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.

Krumhansl, Carol L., and Roger N. Shepard. 1979. “Quantification of the Hierarchy of Tonal Functions within a Diatonic Context.” *Journal of Experimental Psychology: Human Perception and Performance* 5 (4): 579–94. https://doi.org/10.1037/0096-1523.5.4.579.

Lévêque, Yohana, and Daniele Schön. 2015. “Modulation of the Motor Cortex during Singing-Voice Perception.” *Neuropsychologia* 70: 58–63. https://doi.org/10.1016/j.neuropsychologia.2015.02.012.

Marie, Céline, Takako Fujioka, Leland Herrington, and Laurel J. Trainor. 2012. “The High-Voice Superiority Effect in Polyphonic Music is Influenced by Experience: A Comparison of Musicians Who Play Soprano-Range Compared with Bass-Range Instruments.” *Psychomusicology: Music, Mind, and Brain* 22 (2): 97–104. https://doi.org/10.1037/a0030858.

Marie, Céline, and Laurel J. Trainor. 2014. “Early Development of Polyphonic Sound Encoding and the High Voice Superiority Effect.” *Neuropsychologia* 57: 50–58. https://doi.org/10.1016/j.neuropsychologia.2014.02.023.

Marvin, Elizabeth West. 1995. “Research on Tonal Perceptions and Memory: What Implications for Music Theory Pedagogy.” *Journal of Music Theory Pedagogy* 9: 31&nash;70.

——————. 2007. “Absolute Pitch and the Pedagogy of Relative Pitch,” *Journal of Music Theory Pedagogy* 21: 1–34.

——————. 2018. “What I Know Now: Reflections on Music Theory Pedagogy.” In *The Norton Guide to Teaching Music Theory*, ed. Rachel Lumsden and Jeffrey Swinkin, 364–77. W.W. Norton.

Matsunaga, Rie, and Jun-ichi Abe. 2012. “Dynamic Cues in Key Perception.” *International Journal of Psychological Studies* 4 (1): 3–21. https://doi.org/10.5539/ijps.v4n1p3.

Meister, Ingo G., Timo Krings, Henrik Foltys, Babak Boroojerdi, Mareike Müller, Rudolf Töpper, and Armin Thron. 2004. “Playing Piano in the Mind—An fMRI Study on Music Imagery and Performance in Pianists.” *Cognitive Brain Research* 19 (3): 219–28. https://doi.org/10.1016/j.cogbrainres.2003.12.005.

Moreno, Jairo. 2003. “Subjectivity, Interpretation, and Irony in Gottfried Weber's Analysis of Mozart's ‘Dissonance’ Quartet.” *Music Theory Spectrum* 25 (1): 99–120. https://doi.org/10.1525/mts.2003.25.1.99.

Pallesen, Karen Johanne, Elvira Brattico, Christopher J. Bailey, Antti Korvenoja, Juha Koivisto, Albert Gjedde, and Synnöve Carlson. 2010. “Cognitive Control in Auditory Working Memory is Enhanced in Musicians.” *PLOS ONE* 5 (6): e11120, 1–12. https://doi.org/10.1371/journal.pone.0011120.

Pfordresher, Peter Q., Andrea R. Halpern, and Emma B. Greenspon. 2015. “A Mechanism for Sensorimotor Translation in Singing: The Multi-Modal Imagery Association (MMIA) Model.” *Music Perception* 32 (3): 242–53. https://doi.org/10.1525/mp.2015.32.3.242.

Richards, Mark. 2017. “Tonal Ambiguity in Popular Music’s Axis Progressions.” *Music Theory Online* 23 (3). https://doi.org/10.30535/mto.23.3.6.

Ryan, Richard M., and Edward L. Deci. 2000. “Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions.” *Contemporary Educational Psychology* 25 (1): 54–67. https://doi.org/10.1006/ceps.1999.1020.

Saslaw, Janna. 1990. “Gottfried Weber and Multiple Meaning.” *Theoria* 5: 74–103.

Schulze, Katrin, Karsten Müller, and Stefan Koelsch. 2011. “Neural Correlates of Strategy Use During Auditory Working Memory in Musicians and Non-Musicians.” *European Journal of Neuroscience* 33: 189-196. https://doi.org/10.1111/j.1460-9568.2010.07470.x.

Telesco, Paula J. 2013. “Teaching Elementary Aural Skills—How Current Brain Research May Help.” *Journal of Music Theory Pedagogy* 27: 211–46.

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.

Temperley, David, and Elizabeth West Marvin. 2008. “Pitch-Class Distribution and the Identification of Key.” *Music Perception* 25 (3): 193–212. https://doi.org/10.1525/mp.2008.25.3.193.

Trainor, Laurel J., Céline Marie, Ian C. Bruce, and Gavin M. Bidelman. 2014. “Explaining the High Voice Superiority Effect in Polyphonic Music: Evidence from Cortical Evoked Potentials and Peripheral Auditory Models.” *Hearing Research* 308: 60–70. https://doi.org/10.1016/j.heares.2013.07.014.

Footnotes

1. In fact, listeners can also pick up scale-degree frequency distributions in unfamiliar repertoires fairly rapidly and without conscious effort, as Castellano, Bharucha, and Krumhansl (1984) showed with Western listeners exposed to music of North India. Temperley and Marvin (2008) demonstrated that listeners can infer the correct tonic and mode (major or minor) from constructed melodies that maintain the distributions of major- or minor-key profiles, but which present those tones in quasi-random order without schematic patterns of tonality (like $\hat{3}$–$\hat{2}$–$\hat{1}$).
Return to text

2. Karpinski quotes several authors whose discussions touch on key ambiguity: Butler 1989 in paragraph [5] and Matsunaga and Abe 2012 in [13]. In both cases, the described perceptual strategy is for the listener to determine a tonal center quite early in the piece and to keep it in place as a perceptual anchor to which all other tones relate, until another center “defeats it,” to use Butler’s words. I am struck by how similar this idea is to Gottfried Weber’s 1832 “Principle of Inertia” (see Saslaw 1990 and Moreno 2003 for further discussion). Matsunaga and Abe (2012) converge on language similar to Weber’s with their idea of “perceptual inertia.”
Return to text

3. A number of studies have found that musicians show superior auditory working memory and attentional control in comparison with nonmusicians in these areas (e.g., Pallesen et al. 2010; George and Coch 2011; Hansen, Wallentin, and Vuust 2013). We cannot infer a causal effect of music training on these abilities, however, since it may instead be that individuals with high abilities in these areas choose to pursue music training.
Return to text

4. For a comprehensive literature review of auditory imagery (published up to 2010) that includes not only musical imagery, but also imagery for speech and environmental sounds, see Hubbard 2010. Hubbard’s overview goes beyond the scope of our musical topic, but it provides helpful context for the study of auditory imagery in relation to perception and memory and also individual differences between people in their ability to imagine sound.
Return to text

5. The *Journal of Music Theory Pedagogy* has published a number of articles over the years that have brought together music cognition and music theory pedagogy. Among these are Chenette 2018; Marvin 2007, 1995; and Telseco 2013.
Return to text
