Musical Illusions and Phantom Words: How Music and Speech Unlock Mysteries of the Brain. By Diana Deutsch. New York: Oxford University Press, 2019, xviii + 243 pages.

Daniel Shanahan

Music Theory Spectrum, mtaa026, https://doi.org/10.1093/mts/mtaa026
Published: 07 March 2021

"There is no greater impediment to the advancement of knowledge than the ambiguity of words," wrote Thomas Reid (1710–1796), in a fit of wrongness.[1] We can actually learn a great deal about how the mind works through the study of linguistic ambiguity. For example, in Mercer Mayer's *Frog, Where are You?*, a series of drawings conveys a wordless story about a boy realizing his pet frog is missing, setting off to find the frog, and then eventually finding it.[2] This seems like a relatively straightforward narrative, but Mayer's book stands as an important text in the study of the relativity of language. Investigating how speakers of different languages and ages tell this story has taught us a great deal about development (how the narrations change as the readers grow and learn more of their language), filtering (the choices made by the narrator, depending on age and language), and packaging (how parts of the story are compiled into coherent ideas). For example, half of Spanish-speaking three-year-olds mentioned the search for the frog, but only one Hebrew- and one Turkish-speaking child did, and no children who natively spoke English mentioned it at all.[3] Berman and Slobin, who edited a collection of essays studying the frog story, write that:

> We began the study with an expectation that there was a basic set of semantic notions that all children would try to express by some means or other, whether or not grammatically marked in their language … We were repeatedly surprised to discover how closely learners stick to the set of distinctions that they have been given by their language.[4]

The frog story is a tool to examine deeper questions of linguistic relativity; more specifically, it allows us to study how enculturation, learning, and cognition can inform how we interact with utterances that are both narratively ambiguous and semantically consistent. To some extent, the story presents a single succession the events for each reader, although each subject likely makes use of different linguistic resources as they construct narratives.

In *Musical Illusions and Phantom Words: How Music and Speech Unlock Mysteries of the Brain*, Diana Deutsch, for decades a pioneer in the study of how we hear (and mishear) sonic stimuli, similarly presents a number of perceptual phenomena that serve as a window into how we construct order out of ambiguity. In doing so, she provides an insight into the role of development, enculturation, and individual differences in our perception of music. Our responses to the illusions foregrounded in Deutsch's work provide a window into human perception as a whole. To take just one example: a musical tone is composed of many features, including pitch, loudness, timbre, and so on, and Deutsch argues that "we perceive each tone as a bundle of attribute values." Yet she observes that there are instances in which the various components in this perceptual bundle can be at odds with one another, leading them to "recombine incorrectly" (188). Given the right circumstances, we detect "phantom words." We hear impossible continuously ascending or descending tones. We insert melodic fragments into strictly nonmusical speech. These, and many other such anomalies, can tell us a great deal about how our minds combine and parse the attributes of sonic sensation—and what it means to perceive an auditory object in the first place.

Musical illusions are not only good for understanding human perception in general; Deutsch argues convincingly that they also illuminate our understanding of music theory.[5] Illusions—and experiments involving them—are, in essence, instruments of music theory. Alex Rehding, borrowing the concept from Rheinberger's "instruments of science," argues that an instrument of music theory is an object or a process that allows for the exploration or understanding of a music-theoretic phenomenon.[6] (Rheinberger himself focuses on "the power of material objects—in contrast to ideas or concepts—as driving forces in the process of knowledge acquisition."[7]) Viewed through this lens, Deutsch's experiments could be seen as a way of getting closer to an understanding of the "epistemic thing" of music theory, Rheinberger's term for that which is not simply the phenomenon being studied nor the object used to study but rather something that can lead to a deeper understanding. The illusions and experiments presented in *Musical Illusions and Phantom Words*, thus, are able to focus on a host of critical musical and music-theoretical questions: the nature of individual differences between listeners; the role of enculturation and development; how our perception of music is informed by multiple musical parameters; and how we perceive order despite—and sometimes by virtue of—ambiguity. I will consider each in turn.

Musical Illusions and Phantom Words examines issues of individual differences, and the role of enculturation in shaping them, from a number of different angles, each of which contains two layers of interest. The first concerns the existence of the illusion in the first place: why, for example, is it that when an interval is played an octave apart with alternating rests, we hear it as going from ear to ear? The second layer concerns how certain populations perceive can these effects differently than others: why, for example, is this illusion heard one way by right-hand dominant people and another way for southpaws? Each of these questions presents numerous implications for how we frame music analysis, and how we might incorporate such perceptual relativity into our own discussions and practices as analysts.

Deutsch provides a brief intellectual history of the study of brain localization, which is appreciated and necessary, as much of this early research is intertwined with problematic—and very often racist—pseudoscience. Franz Joseph Gall (1758–1828) first argued that the mind was an aggregate of specific physical spots in the brain (with specialized "organs" for love, the tendency to commit crimes, courage, etc.). In seeking to support his theory, Gall collected skulls of those known to exhibit each trait. Johann Spurzheim (1776–1832), a student of Gall, would coin the term phrenology, and together they would become quite respected scholars in the early nineteenth century. Eventually, the misguided nature of phrenology would become self-evident, and it faded from serious conversation as the science of neuroanatomy matured. Scientists continued to examine patients with brain injuries as a way of understanding localization in the brain. It is refreshing that Deutsch takes the time to discuss the problematic origins of cortical localization and the direct lines of inquiry that would lead to Paul Broca's influential (and eugenicist) work of the nineteenth century, which would lead in turn to the common acceptance of the notion of the functionally localized—and hemispheric—brain.

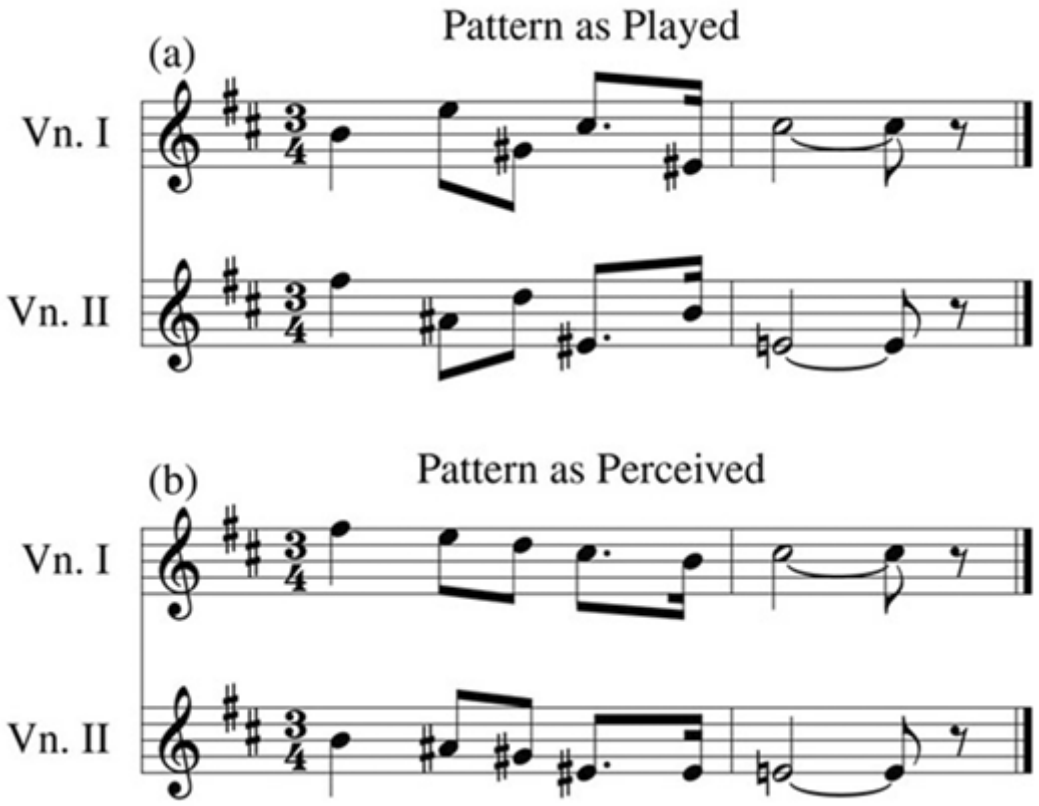
After discussing localization and handedness, Deutsch presents her work on the octave illusion in an almost auto-ethnographic style. In this study, two tones are presented an octave apart in both ears, but the listener hears two separate melodic sequences. Deutsch writes that "my right ear heard 'high tone–silence–high tone–silence' while at the same time my left ear heard 'silence–low tone–silence low tone'" (24). The experiment begins as an illusory conjunction: multiple stimuli are being combined into one perceived object. That the octave illusion seems to be related to handedness—which ear is hearing the high tone or the low tone seems to change for left-handers—demonstrates that the "correct" way to perceive the tones, and their separation into independently perceived melodic lines, varies greatly between individuals but is nevertheless constrained by the handedness of the individual.

A similar effect can be seen in the glissando illusion, in which a synthesized oboe tone at a fixed pitch alternates between left and right stereophonic channels, while a sine-wave glissando simultaneously alternates in the opposite channels, right to left and back. That different subjects perceive this

phenomenon in different ways is interesting in itself, but the fact that those differences can broadly be correlated with handedness is fascinating. Right-handed individuals tend to hear the glissando component not as a discrete slice of pitch information but as a seamless, stereophonically directed continuum, gradually traveling from left to right as the pitch ascends and then back to the left ear as it descends. Left-handed listeners, however, do not consistently hear this spatialization, either right to left or left to right. Additionally, right-handed listeners tend to associate a *vertical* directionality to the stimulus more consistently (lower pitches are heard as closer to the ground). Right-handers are more likely to associate higher pitches on their right, lower tones on their left, and they are more able to accurately identify tone combinations when they conform to this cross-modal mapping (42). Here, Deutsch points to a disconnect between what might be best perceptually and what is common practice: symphony orchestras are arranged with the higher instruments on the left and the lower instruments on the right, adhering to this mapping from the perspective of the performers, rather than the perspective of the audience. Although there aren't many practical solutions for live music (performers either wouldn't be able to hear each other too well, the orchestra probably shouldn't turn their backs to the audience, etc.), Deutsch does argue for flipping one's speakers around when listening to orchestral music at home.

The scale illusion is another example where subjects perceive coherence and continuity despite ambiguity. In this case, two highly disjunct melodies produced by separate instruments are heard as forming a single, uninterrupted scale. The listener is able to create a coherent melodic idea when none is actually sonically present. Here, Deutsch discusses the (somewhat apocryphal) story of a rift between Tchaikovsky and Artur Nikisch concerning the orchestration of the *Pathétique* Symphony, specifically the section of the fourth movement in which theme and the accompaniment alternate between the two violin parts, reproduced in Example 1.[8] Deutsch writes that Nikisch's vexation with this scoring would lead to his own reorchestration and a schism in performance practice of the piece, each derived from the composer and the conductor having heard the separation of these lines differently. This anecdote suggests ways in which musical illusions are not just idle, contrived curiosities but auditory realities that might be exploited for aesthetic purposes.

EXAMPLE 1:



Deutsch, Figure 2.6 (The beginning of the final movement of Tchaikovsky's Sixth Symphony—*The Pathétique*). Originally printed in *Musical Illusions and Phantom Words: How Music and Speech Unlock Mysteries of the Brain* (Deutsch 2019, 36). Reproduced with the permission of Oxford University Press through PLSclear.

The tritone paradox, discovered by psychologist Robert Shepard in the 1960s, provides another window through which we might begin to discuss why perceptual disparities across individuals exist and what biological and/or cultural individual differences might account for them. When presented with a Shepard tone followed by another a tritone away, some participants would hear the tones as ascending (e.g., C4–F♯4), whereas others would hear it as descending (e.g., C4–F♯3) (71–81).[9] For Deutsch, a British-born resident of San Diego, this disparity was originally quite surprising, as the majority of her American students heard the tones moving in the *opposite* direction as she had heard. After examining possible factors related to musical training, hearing ability, and presenting the stimuli in separate ears, it appeared that the difference in perception resulted from a difference in dialects. Put succinctly, one's native dialect dictates which intervals and pitch ranges a speaker is most accustomed to. There is also a hypothesized evolutionary role for dialects, and our understanding and perception of them, based on pitch content and range. There is an adaptive advantage in protecting family, and even risking one's own life to protect a family member, and similar dialects might, in turn, mean closer familial relationships. Deutsch argues that the awareness of the intricacy of dialects through exposure becomes ingrained in one's memory. The tritone paradox, like many of these illusions, demonstrates the effects of memory and expectation in how we perceive and organize musical patterns.

In order to further explore this hypothesis, Deutsch followed this study up with a large-scale experiment carried out in collaboration with BBC Radio; hundreds of participants supported these findings: listeners who had been raised in the south of England heard the sequence of tones in the same way Deutsch had. Subsequent research would compare how Vietnamese immigrants and children of immigrants in California responded to the tritone paradox; it found that both Vietnamese groups perceived the direction of melodic motion differently from the native Californians but did not differ from one another. Specifically, the pitches between E and A were more likely to be heard as descending

by the Vietnamese groups, and the pitches between B ♭ and E ♭ were more likely to be heard as descending by the participants from California. Deutsch argues that "these results strongly support the view that perception of this illusion reflects a speech-related template that develops early in life, and survives into adulthood" (77). As with all the paradoxical auditory phenomena discussed in the book, Deutsch invites the reader to judge the effects for themselves—QR codes linking to online versions of each illusion are conveniently placed throughout. It is surprising how strong the effect of these experiments can be, even when conducted informally and even when fully aware intellectually of how they work.

The connection to linguistic exposure is also present in Deutsch's discussion of absolute pitch. There have been a number of studies examining the prevalence of absolute pitch, which, depending on how absolute pitch is defined, is far more prevalent than we might think. Listeners are able to discriminate between at-pitch and pitch-shifted versions of well-learned sonic stimuli, including the dial tones of landline telephones, the tone that television censors use to beep out swear words, and excerpts of their favorite songs and segments from film and television soundtracks.[10] For those whose native language is tonal, pitch is associated very early on with words spoken and heard. As Deutsch notes:

> … brain circuitry for absolute pitch is already in pitch, so [native speakers of tonal languages] develop absolute pitch for musical tones similarly to the way they would acquire words in a different tone language. In contrast, nontone languages such as English use pitch for purposes such as conveying grammatical structure, emotional tone, and other aspects of prosody, but not for determining the meaning of individual words. In consequence, speakers of non-tone languages are at a disadvantage for acquiring absolute pitch. (89)

To test the linkage of absolute pitch and spoken language, Deutsch and her colleagues worked with native Vietnamese speakers with no musical training, asking them to read ten words in Vietnamese, repeated over the course of multiple days. Speakers were remarkably consistent in maintenance of lexical pitch—often within a semitone. Additional studies examined both the role of musical training and speaking a tone-language natively. Average scores for absolute pitch were higher for those who received musical training earlier, but the largest effect came from fluency with a tone language. This linkage led Deutsch to conclude that "there is a strong relationship between the possession of absolute pitch and the language spoken by the listener—the prevalence of this ability is far higher among tone language speakers than among speakers of non-tone language such as English" (102). Although the study of the prevalence of absolute pitch might not have readily apparent applications to music theory, the "epistemic thing" being examined is something at the heart of many music-theoretic questions: the relationship between prosodic and melodic structures and the role enculturation plays in the understanding of these structures. As theorists strive to decenter the way in which they analyze music and studies have more frequently explored the role of tone languages on melodic construction from a music-theoretical point of view, it is all the more important to engage with research such as this, which examines the role of linguistic and cultural differences on melodic and prosodic perception.[11]

In Chapter 10, Deutsch discusses the well-known speech-to-song illusion, in which the phrase "sometimes behaves so strangely" is repeated until it sounds increasingly melodic. This, too, can tell us a great deal about enculturation and learning; those that have heard the fragment repeated to the point of semantic satiation will hear the piece to be musical, whereas those who have not been exposed to the repetition will hear it as speech. Deutsch provides a bit of background to the experiment here, discussing how she had been preparing an introductory spoken commentary to a CD and had been listening to phrases on a loop in order to find any errors in the recorded speech. She mentions how she put one phrase ("sometimes behave so strangely") on a loop and began working on something else entirely, only to be drawn back into it by the sheer musicality of the loop. This leads to a discussion of the relatively fluid distinction between speech and sound, touching on aspects of *Sprechstimme* and opera, as well as the overlap of cognitive processing that occurs in both speech and music, reminiscent

of Patel's "shared structural integration resource" hypothesis, which argues that similar cognitive processing resources are utilized by both speech and music.[12] It is interesting to track the life of this illusion as it has been built upon so much since its introduction. For example, Simchy-Gross and Margulis have recently shown how nonmusical, non*linguistic* sounds can similarly be perceived as increasingly musical with repetition. But repetition alone is not enough: what makes something sound melodic is both the repeated exposure and the realization that the sound is in fact musical, and that once heard as such, it is very difficult to unlearn the musicalized connotations that come with that exposure.[13] There are important implications from these auditory phenomena that must be heeded by music theorists: not only is the line between speech, noise, and melody fluid, but what even counts as a music "thing" is not a fixed category but something heavily dependent upon listener exposure.

The illusions presented by Deutsch demonstrate how a musical idea might be inferred and understood despite its ambiguous nature. We hear continuously ascending or descending Shepard tones as never-ending if they are constructed properly, and we hear phantom sounds and words, such as when records are played in reverse, where there are none (65–66, 103–115). I would argue that the "epistemic thing" is both the musical order that we infer *and* the nature of how our inferences are consistently varied based upon certain constraints, such as the exposure to languages and dialects, handedness, and so on. Much like Leonard Meyer's work of the 1950s, Deutsch's theorization rests heavily on the principles of Gestalt psychology—continuity, grouping, and coherence.[14] She spends much of Chapter 3 discussing some of the work of scholars who have focused on the influence of multiple parameters on the perception of the musical whole. For example, Van Noorden's demonstration of "the galloping rhythm" between voices, which demonstrates how pitch can affect our perception of rhythm as a whole, Dowling's study of how melodic independence is facilitated by octave differentiation or Wessel's timbre illusion, which demonstrated the continuous nature of how, when timbral differentiation is significant, multiple melodic lines are heard, but as the sounds grow closer in pitch space, they are more likely to be heard as a single melodic line.[15]

Musical Illusions and Phantom Words is the culminating product of a brilliant and trailblazing figure who for decades has been at the forefront of music psychology. Deutsch compiles decades' worth of findings to provide a cohesive narrative about how the study of the sonically ambiguous can lead to a greater understanding of both the mind and music. Her work pushes the reader to avoid easy assumptions, to examine multiple rather than singular musical parameters, to steer away from the dualistic trap of "nature vs. nurture." In highlighting how the perceptual organization of musical illusions is neither simply universal nor entirely individualized, Deutsch is pushing us to study the deeper points of musical relativity. *Musical Illusions* is a crucial reminder to music theorists that the framing and understanding of a musical object are never static, always culturally situated, multivalenced, and complex.

Footnotes

1 Reid (1983, 129).

2 Mayer (1969). I am indebted to Larry Zbikowski, who pointed me in the direction of the frog story corpus research.

3 See Berman and Slobin (2013). See also Gerrig (1995) for a review of this work.

4 Ibid (614). See also Gerrig (1995).

5 After acknowledging the work of Galilei, Mersenne, and Descartes, Deutsch writes that "I strongly believe, and hope to convince you also, that the time has come to place music theory on a firm experimental footing" (8). This firm footing is, according to Deutsch, made possible by our ability to computationally generate experimental stimuli that can exploit perceptual ambiguities.

6 Rheinberger (1997) and Rehding (2016).

7 Rehding (Ibid, §4.1).

8 Deutsch cites a discussion of her work in *Scientific American* on the topic written by Shawn Carlson (1996, 112–15). It is unclear why Tchaikovsky, who called Nikisch "a conductor of genius" would be so adamantly opposed to the change, but it is nevertheless an interesting discussion point that relates performance history to music perception. As David Brown (1992, 455) observes, Tchaikovsky had used this "seesaw" effect before in m.74 of *The Storm* and in m.142 of the Third Suite (*Valse mélancolique*). Deutsch's reference to the effects of seating arrangements of the orchestra in the nineteenth century might also be problematized a bit more, as Daniel Koury (1986) points out that we don't really have the seating plans of nineteenth-century Russian orchestras. On this subject, see also Deutsch (2013). Many thanks to Joseph Kraus for discussing this point with me.

9 A Shepard tone is a synthetic note generated from a combination of sine waves of the same pitch class in multiple octaves, with mid-frequency components having the highest amplitude and low- and high-frequency components lower amplitude. A Shepard scale takes a Shepard tone and moves it, either discretely or continuously, up or down in frequency onto new Shepard tones centered on different pitch classes. The continuously shifting glissando form is referred to as the "Risset glide" after French composer and physicist Jean-Claude Risset.

10 See Smith and Schmuckler (2008), Van Hedger, Heald, and Nusbaum (2016), Levitin (1994), and Schellenberg and Trehub (2003).

11 See, for example, Aaron and Quintina Carter-Énji's (2020) work on Yorubá.

12 See Patel (1998).

13 See Simchy-Gross and Margulis (2018).

14 See, for example, Meyer (1956).

15 See van Noorden (1975), Dowling (1973), and Wessel (1979).

Works cited

Berman Ruth A., Slobin Dan Isaac. 2013. *Relating Events in Narrative: A Crosslinguistic Developmental Study*. Abington: Psychology Press.
Google Scholar

Brown David. 1992. *Tchaikovsky:(1885–1893). The Final Years*. New York and London: W. W. Norton.
Google Scholar

Carlson Shawn. 1996. "Dissecting the Brain with Sound." *Scientific American* 275 (6): 112–15.
Google Scholar

Carter-Énjì Aaron, Carter-Énjì Quintina. 2020. " 'Bold and Ragged': A Cross-Cultural Case for the Aesthetics of Melodic Angularity." *Music & Science* 3: 1–15.
Google Scholar

Dowling W. Jay. 1973. "The Perception of Interleaved Melodies." *Cognitive Psychology* 5 (3): 322–37.
Google Scholar

Gerrig Richard J. 1995. "Review of Relating Events in Narrative: A Crosslinguistic Developmental Study by Ruth A. Berman and Dan Isaac Slobin." *Language* 71 (4): 806–08.
Google Scholar

Koury Daniel J. 1986. *Orchestral Performance Practices in the Nineteenth Century: Size, Proportions, and Seating*. Rochester: University of Rochester Press.
Google Scholar

Levitin Daniel. J. 1994. “Absolute Memory for Musical Pitch: Evidence from the Production of Learned Melodies.” *Perception & Psychophysics* 56 (4): 414–23.
Google Scholar

Mayer Mercer. 1969. *Frog, Where Are You?* New York: Dial Press New York.
Google Scholar

Meyer Leonard B. 1956. *Emotion and Meaning in Music*. Chicago: University of Chicago Press.
Google Scholar

Noorden Leo Paulus Antonie Servatius van. 1975. *Temporal Coherence in the Perception of Tone Sequences*. Vol. 3. Netherlands: Institute for Perceptual Research Eindhoven.
Google Scholar

Patel Aniruddh D. 1998. “Syntactic Processing in Language and Music: Different Cognitive Operations, Similar Neural Resources?” *Music Perception: An Interdisciplinary Journal* 16 (1): 27–42.
Google Scholar

Reid Thomas. 1983. *Thomas Reid’s Inquiry and Essays*. Eds. Beanblossom Ronald and Lehrer Keith. Indianapolis: Hackett Publishing.
Google Scholar

Rehding Alexander. 2016. “Instruments of Music Theory.” *Music Theory Online* 22 (4).
https://mtosmt.org/issues/mto.16.22.4/mto.16.22.4.rehding.html.
Google Scholar

Rheinberger Hans-Jorg. 1997. *Toward a History of Epistemic Things: Synthesizing Proteins in the Test Tube*. Stanford, California: Stanford University Press.
Google Scholar

Schellenberg E. Glenn, Trehub Sandra E.. 2003. “Good Pitch Memory Is Widespread.” *Psychological Science* 14 (3): 262–66.
Google Scholar

Simchy-Gross Rhimmon, and Margulis Elizabeth H.. 2018. “The sound-to-music illusion.” *Music & Science* 1:1–6.
https://doi.org/10.1177/2059204317731992.

Smith Nicholas A., Schmuckler Mark A.. 2008. “Dial A440 for Absolute Pitch: Absolute Pitch Memory by Non-Absolute Pitch Possessors.” *The Journal of the Acoustical Society of America* 12 (4): 77–84.
Google Scholar

Van Hedger Stephen C., Heald Shannon L. M., Nusbaum Howard C.. 2016. “What the [Bleep]? Enhanced Absolute Pitch Memory for a 1000 Hz Sine Tone.” *Cognition* 154: 139–50.
Google Scholar

Wessel David L. 1979. “Timbre Space as a Musical Control Structure.” *Computer Music Journal* 3 (2): 45–52.
Google Scholar