



## Looking Beyond the Score: The Musical Role of Percussionists' Ancillary Gestures

Michael Schutz and Fiona Manning

NOTE: The examples for the (text-only) PDF version of this item are available online at:

[http://www.mtosmt.org/issues/mto.12.18.1/mto.12.18.1.schutz\\_manning.php](http://www.mtosmt.org/issues/mto.12.18.1/mto.12.18.1.schutz_manning.php)

KEYWORDS: music cognition, perception, visual information, gestures, performance, sensorimotor integration, duration

ABSTRACT: Performing musicians frequently use physical gestures that are more elaborate than required for sound production alone. Such movements are not prescribed in traditional musical scores, nor are they evident in audio recordings, and consequently they are rarely regarded as integral to a formal musical analysis. However, there is growing evidence that these movements do in fact alter an audience's listening experience—i.e., the way a performance “sounds.” Therefore, we believe that analyses of these movements can inform more traditional analyses of notes and rhythms by lending insight into the way in which these musical elements are *perceived*. Here, we review research on the role of gestures in shaping the musical experience, focusing in particular on gestures used by percussionists to control perceived note duration. This paper embraces the multi-media affordances of *Music Theory Online* by integrating stimuli from key experiments—the first publication of these materials. Our aim is not only to summarize a growing body of work on the musical role of extra-acoustic factors such as ancillary gestures, but also to present new avenues of musical research that complement existing approaches.

*Received July 2011*

---

[1] The written score frequently serves as the basis for our efforts to understand music's structure, content and meaning. Although scores capture many important aspects of a composition, certain elements are difficult to analyze and impossible to understand without accounting for the way in which the acoustic signal (represented abstractly by the score) is *perceived*. For example, although they share identical intervallic relationships, our perception of the two triads in **Example 1** differs markedly. The higher one sounds “sweet” and “pure” whereas the lower sounds “rough” and “muddy.” This appreciably different listening experience is attributable not to the structure of the triads themselves, but rather the structure of the perceptual system—in particular the differential sizes of critical bands in low vs. high frequency ranges.<sup>(1)</sup> This is but one example of how information beyond the score (in this case the listening apparatus used to experience sound) shapes music listening. Other factors are subtle and more difficult to recognize, yet also play a crucial role. For example, the next section summarizes ways in which a performer's body movements routinely shape the listening experience. Given mounting

evidence documenting the musical importance of gesture, we believe that efforts to understand music can and will benefit from exploring this perspective.

[2] Our goal in this paper is to build a case that (1) gestures used by performers play a meaningful role in music perception even though they are not represented in the score and (in some cases) do not have acoustic consequences, (2) new research tools coupled with the traditional techniques and methodology of music perception allow for precise analysis of these gestures (with a degree of rigor traditionally reserved for “the notes on the page”), and (3) much as the thoughtful analysis of a score can be insightful in understanding a musical composition, analysis of a performer’s body movements can also be informative. Although much of this article will focus on one particular class of gestures (those used by percussionists), this issue is broadly relevant for all musicians (in addition to artists, scholars, and critics). Therefore before delving into a focused review, we will discuss the relationship between gesture and music broadly, as well as commonalities between the use of gesture in music and dance. The second section will then focus on one particular type of ancillary gesture used to overcome acoustic limitations of the marimba. The third and fourth sections will explore the perceptual basis for this phenomenon, and the fifth will review efforts to analyze and deconstruct the gestures themselves. The final section will discuss future directions for this line of work, as well as general implications of these approaches for musical research.

## 1. The Use of Gesture in Music

[3] Interest in the role of gestures in music is a vast topic, one that has seen significant research attention in recent years. The term “gesture” itself can be interpreted in multiple ways. At a basic level, it may refer either to a particular segment of music (see [Hatten 2004](#)) or to a physical motion used by performers. The focus of this paper is on the latter. Wanderley and colleagues ([2005](#)) distinguish further between two classes of physical gestures—*effective gestures* that are required for sound production, and *ancillary gestures*, which are not necessary for the creation of sound. *Ancillary gestures* have been previously referred to as either expressive movements ([Davidson 1993](#)) or body language ([Dahl and Friberg 2007](#)). These movements are often thought to be of secondary importance, given that they lack significant acoustic consequences and their production is rarely consciously/actively regulated by performers. However, there is significant stability in these gestures across performances by a single musician ([Wanderley 2002](#)); and as discussed in the next subsection, they are capable of systematically affecting an audiences’ listening experience.

### 1.1 Ancillary Gestures in Music

[4] Ancillary gestures can play a profound role in music listening—despite their lack of acoustic consequences. For example, judgments of tension and phrasing in the second of Stravinsky’s *Three Pieces for Clarinet Solo* differed significantly when participants watched the performer rather than listened to the audio alone ([Vines et al. 2006](#)). This phenomenon is not limited to clarinetists—emotions including happiness, sadness and anger can be readily communicated through gestures on a number of instruments ([Dahl and Friberg 2007](#)). In fact, in some cases these distinctions may be more clearly discerned through a visual, rather than an auditory presentation ([Davidson 1993](#)). This significant role of vision carries important implications for the evaluation of performers—seeing the body movements of marimbists playing in an expressive style can affect ratings of audience interest ([Broughton and Stevens 2009](#)).

[5] The role of visual information is limited not only to judgments of high-level characteristics such as musical expression or performance quality, but also extends to low-level characteristics such as judgments of pitch intervals ([Thompson et al. 2005](#)), pitch accuracy ([Gillespie 1997](#)), and loudness ([Rosenblum and Fowler 1991](#)). For example, seeing a cellist’s bowing and plucking motions affects perceptual ratings of the concurrent note’s timbre—i.e., bowed vs. plucked ([Saldaña and Rosenblum 1993](#)). Therefore the visual information associated with a performer’s physical gestures plays an important role in shaping the musical experience on several levels (for a comprehensive review see [Schutz 2008](#)).

### 1.2 Ancillary Gestures *as* Music

[6] The twentieth century has seen the role of gesture in music reach new heights—particularly in music for percussion. For example, in the marimba solo *Six Elegies Dancing* ([1987](#)), composer Jennifer Stasack gives elaborate instructions on the

motions to be used while performing—many of which have no acoustic consequences. The preponderance of compositions emphasizing gestures have led to a sub-genre of “theatrical percussion” capitalizing on the tight relationship between gestures, music for percussion, and perception (due to the large amount of physical motion required to play such instruments, music for percussion is a particularly fertile ground for exploring such connections).<sup>(2)</sup>

[7] Although it is gaining popularity in new music, this focus on gesture is far from “new.” John Cage uses gestures to great effect in a number of compositions such as *Living Room Music* (1976), which combines elements of percussion and theatre. In this composition, performers are called upon to play rhythms on “found objects” such as cups, bowls, books, and other items commonly situated in a living room. The variety of creative realizations of this score demonstrates the integral role of body movements, as performers frequently add gestures for reasons as much theatric as acoustic. The use of ancillary gestures as music raises interesting questions about links with another form of expression built upon the use of movement over time—dance.

### 1.3 Ancillary Gestures as Dance

[8] Like ancillary gestures in music, dance frequently occurs concurrently with acoustic information. In fact, some have gone so far as to define dance as “human movement that is formalized...to the accompaniment of music or other rhythmic sounds...” (Van Camp 1981). In a sense, musicians “dance” when performing, given that their ancillary movements accompany the music without affecting its acoustic characteristics. Therefore, one way of viewing dance is essentially as a special case of ancillary gestures in that it is a series of movements accompanying music not required for sound production.

[9] Dance choreography frequently reflects musical structure as dancers’ movements are generally designed to accompany and interact with concurrent musical events. Consequently, ratings of section breaks, tension, and emotion when viewing ballet show a strong similarity for the separate music and dance components (Krumhansl and Schenck 1997). Similarly, observers consistently rate dance as more realistic, creative and natural when the performance is generated to match the relative changes in musical features (Kim et al. 2009). These findings illustrate that movement can be used to communicate structural features of a musical composition, with particular movements conveying specific musical characteristics.

[10] Dance movements made in response to music listening (i.e., “music-induced movement”) not only react to the low-level temporal structure of music, but also reflect the rich hierarchy of temporal information by which it is organized. For example, when asked to move freely while listening, movement of the extremities tends to synchronize with faster metric levels whereas movement of the torso tends to synchronize with slower metric levels (Toiviainen et al. 2010). Therefore, analysis of dance in this context actually informs our understanding of musical structure as it lends insight into the way timing information is hierarchically organized in the minds of listeners.

[11] The relationship between the temporal structure of music and dance is so prominent that some have hypothesized they may have originated as a single system of communication (Hagen and Bryant 2003). Although we may never fully understand the evolutionary development of these domains, broad similarities between the two are consistent with the possibility of a common origin. In fact in many non-Western cultures, music and dance are intertwined into a single, multimodal experience. Although the two often remain distinct within the Western classical tradition, they are presented in tandem in artistic contexts such as ballet, opera and musical theatre. Additionally, the two frequently co-occur in popular music concerts and music videos, which often include elaborately staged sequences of movement in conjunction with the sound.

## 2. Ancillary Gestures as Performance Tools

[12] In addition to reflecting large-scale aspects of structure, ancillary gestures can also be useful in overcoming acoustic limitations of certain musical instruments. In other words, they can be used to accomplish *perceptually* that which is impossible *acoustically*. This is illustrated in a “musical illusion” resolving a long-standing debate amongst percussionists (Schutz 2009) concerning the relationship between the length of the physical gesture used to strike a marimba and the duration of the consequent note.

[13] Longtime New York Philharmonic percussionist Elden “Buster” Bailey observed that, “[When] sharp wrist motions are

used the only possible results can be sounds of a staccato nature... [When] smoother, relaxed wrist motions are used, the player will then be able to feel and project a smoother, more legato-like style” (1963). Others, such as Leigh Howard Stevens, are adamant that gesture length is irrelevant, arguing it has “no more to do with [the] duration of bar ring than the sound of a car crashing is dependent on how long a road trip was taken before the accident” (personal communication, 2004). On the surface, there appears to be merit to each of these competing points of view. A longer swing of the bat intuitively sends a ball farther. However, from the physicist’s perspective, motion after impact is not directly relevant to the acoustic consequences of the preceding event.

## 2.1 Gestures are Acoustically Inconsequential, but Perceptually Useful

[14] To explore this issue, renowned marimbist Michael Burritt (now Professor of Percussion at the Eastman School of Music) performed a series of “long” and “short” notes in a recital hall on a professional-grade marimba (see samples of these gestures in **Example 2**). An analysis of the acoustic information produced by these long and short gestures found no meaningful distinction between the two sounds. Therefore, his long and short gestures failed to create notes that were acoustically distinguishable. However, the following experimental research demonstrates that the failure of this gesture *acoustically* does not necessarily prohibit its success *perceptually*.

[15] To explore this issue, participants were presented with four classes of audio-visual pairings. Two of these were the original long and short gesture-tone pairings produced by the performer, and two of these were “hybrid pairings,” consisting of the long gesture paired with the acoustic signal created by the short gesture and vice-versa. Participants were informed that in some instances the auditory and visual components of the stimuli would be inconsistent, and were asked to judge the duration of the acoustic sound alone (i.e., ignore the visible gesture). Prior to the experiment they performed a warm-up phase in which they were presented with a variety of gesture-tone pairings to familiarize themselves with the range of long and short sights and sounds they would experience in the actual experiment, and to acquaint themselves with the experimental procedure. (3)

[16] Consistent with the previously described acoustic analysis, the auditory components of the stimuli were indistinguishable—ratings of the long gesture paired with the sounds produced by the long and short gestures did not differ (**Example 3**). However, ratings of a sound when paired with the producing gesture differed significantly. In other words, the same sound *sounded* longer when paired with a long gesture than when paired with a short gesture (Schutz and Lipscomb 2007). The fact that this difference occurred despite explicit instructions to ignore visual information suggests that the visible gestures actually altered the participants’ experience of the note. In other words, although the performer’s gesture failed to change the (acoustic) sound of the note, it successfully changed *the way the note sounds* (note: see sections three and four for evidence that this perceptual change is obligatory and pre-conscious).

[17] Given that it demonstrates a clear role of visual information in the perception of music, this result raises interesting questions about what music “is.” Furthermore, it illustrates challenges with attempting to fully understand the musical experience based solely on an analysis of a notated score (or even a CD, which robs a marimbist of the ability to sculpt his or her audiences’ perceptual experience of note duration). The durations of notes perceived by the audience cannot be fully understood merely by examining their notated values in the score, as skillful performers are able to use different kinds of gestures to communicate the composer’s intentions—and their interpretative role by definition involves movements that are “beyond” the notation itself. These results also raise important questions about the relationship between sound (i.e., an acoustic signal) and the way that sound is perceived in the mind of a listener—questions that are crucial to understanding the nature of the musical experience.

## 3. Perceiving the World

[18] The seemingly effortless nature of the perceptual system masks its true complexity. Our ability to perceive and understand the world around us is actually the end result of a complex and fascinating chain of events. In order to organize the many sensations processed concurrently, our perceptual system must frequently make implicit assumptions (i.e., automatically and outside our conscious awareness) about this information so as to present us with an internal *perceptual*

*experience* corresponding to the external *state of the world*. Although errant perceptions (i.e., “illusions”) are intriguing, such errors are really the exception, rather than the norm. Perception involves more than the mere detection of information from the environment. Instead, it involves a complex interaction between events (e.g., a sound) and the implicit assumptions used to process them—assumptions that are crucial both to our everyday perceiving in general as well as music listening in particular.

[19] One instance of a useful assumption is displayed in **Example 4**. Although all of the shaded circles in this figure are identical in their pixel-by-pixel relationships, some are shaded on top and others on the bottom (i.e., some circles are 180 degree rotations of the others). This ambiguity presents a dilemma for the perceptual system, as it is not possible to interpret the three dimensional structure of the object (i.e., its “depth”) based solely on the pattern of light entering the eye. (4) Although the optical information is ambiguous, we consistently and vividly perceive the image with shading at the bottom as a convex “bump” and the image with shading at the top as a concave “dent”. This is due to the (generally correct) implicit assumption that light is coming from above our heads. This implicit knowledge allows the perceptual system to “decode” the otherwise ambiguous pattern of light, giving us a perceptual experience aligned with the physical object (Kleffner and Ramachandran 1992). Because the perceptual system is tuned to our environment, we are able to benefit from implicit assumptions or perceptual short-cuts. By “thinking for us,” our brains provide a perceptual experience that is in fact superior to the information encoded in the light entering our eyes.

[20] These perceptual assumptions are not restricted to a single modality such as vision, but also occur cross-modally when we process concurrent sights and sounds. Our minds work to organize sensory inputs in a meaningful way, automatically binding sensory information together if it appears to originate from the same event. This can be seen clearly in the *McGurk effect* (Video Example 5) in which visual information categorically changes our perception of concurrent speech (McGurk and MacDonald 1976). This phenomenon entails pairing one speech sound with the lip movements used to produce a different sound. The resulting percept is one that is intermediate between the visual and auditory information. In the example below, the lip movements displayed are those used to produce the sound /ga/, whereas the auditory information is actually /ba/. The resultant perception is that of /da/—essentially the acoustic “average” the two. This is a powerful demonstration of the process of sensory binding leading to a unified, multimodal experience.

### 3.1 How Does the Perceptual System “Know” When to Integrate?

[21] Clearly, our perception of information in the environment is influenced by factors beyond the information itself. Implicit assumptions about both the structure of our environment (Example 4) and the multi-modal nature of our experience of that environment (Example 5) play significant, though subconscious, roles in the perceptual process. It is important to note that multi-modal integration is not limited to contrived scenarios with artificially paired sights and sounds (as is the case with the McGurk effect). In everyday perceiving we often experience a number of events occurring simultaneously—such as the cacophony of speech heard at a cocktail party. Audio-visual integration is a constant background process assisting with the organization of a chaotic stream of sights and sounds into the coherent perceptual experience of unified multi-modal events (Kubovy and Schutz 2010).

[22] Consequently, one of the challenges for the perceptual system is to identify sights and sounds originating from common sources. And one of the cues for discerning multi-modal relationships is *causality*. Much as the McGurk effect exploits this cue by pairing spoken sounds with lip movements that plausibly caused them, musicians can take advantage of this same process by using gestures to shape the perception of notes that they perform.

## 4. Causality Cues Integration

[23] The importance of causality in audio-visual integration is illustrated best by viewing the perceptual ramifications of its absence. Manipulations weakening the causal link diminish the strength of the illusion; manipulations breaking it destroy it entirely. For example, sounds that could not be caused by impact gestures such as those of a clarinet or human voice (Example 6) fail to integrate with impact motions. When asked to judge the duration of these sounds independently from concurrent gestures (the same instructions used in the previous experiment), participant ratings indicate that the gestures fail

to influence judgments of auditory duration. This suggests that the perceptual system does not integrate sights and sounds when they do not specify a common event (Schutz and Kubovy 2009a).

[24] That is not to say *any* alternative sounds will fail to bind, as other sounds caused by impact events do in fact integrate. For example, the same gestures paired with the sound produced by a piano (Example 7, left) integrated, and consequently influenced their perceived duration. However, the magnitude of the gestures' effect on the perceived duration of the piano sounds was about half of the magnitude of the effect on the marimba sounds (Example 7, right). This experiment demonstrates that the illusion is contingent upon detection of congruity between the visual motion and the auditory timbre.

[25] The importance of cross-modal causality can also be seen in a second experiment manipulating audio-visual temporal synchrony. In this experiment, participants experienced three kinds of audio-visual pairings. In the first condition, the note occurred slightly before the visual moment of impact (Example 8, left), generating videos in which the sound appeared to “lead” the gesture. In the second condition, the note occurred slightly after the visual moment of impact (Example 8, right), creating the appearance of the sound “lagging” the gesture. The third condition used videos from the original experiment—in which the sight and sound occurred synchronously (participants experienced the videos in a randomized order).

[26] Here, the gestures integrated with the marimba sound in the audio-lag condition, but not the audio-lead condition. This demonstrates the importance of causality in a different context. As the speed of sound is significantly less than that of light, we often experience situations in which sound arrives at our ears slightly later than images arrive at our eyes. Therefore, disruptions to temporal alignment have asymmetrical perceptual consequences—sounds that lag a visual event are integrated preferentially over sounds that lead. Although preserved, the illusion was notably weaker in the audio-lag case than in the synchrony case, which indicates that introducing an auditory lag weakens the causal link (and an audio-lead destroys it completely).

[27] Together, these experiments demonstrate that the illusion is contingent upon cross-modal causality. Manipulations that break the causal link (such as the introduction of non-percussive timbres in the first experiment or temporal ordering inconsistent with the physical world in the second) destroy the illusion. Moreover, the strength of the illusion is related to the strength of this link. Manipulations that weaken (but do not destroy) the causal link weaken the illusion. For example, the gestures integrate with the sound of a piano (which is an impact sound, simply of different materials than depicted in the video). Similarly the gestures integrate with sounds that lag the observed moment of impact—although in both these cases the magnitude of the illusion is less than in the original. This insight is significant as it strongly suggests that the gestures do not merely change the way participants are responding to questions about note duration, they actually change *the way the notes sound*.

[28] Furthermore, these experimental manipulations also highlight that the illusion is robust with respect to prior knowledge. All participants were informed a-priori that the auditory and visual information were in some cases unrelated, and were asked to ignore the visual information when judging auditory duration. They were actually quite adept at doing just that—however, only when the gesture lacked a causal link with the sound. This strongly suggests that the visual influence is obligatory, and that prior knowledge about the illusion does not dampen its salience any more than understanding the McGurk effect (Example 5) weakens the influence of the lip movements on our perception of the spoken syllables.

## 5. Analysis and Deconstruction of the Gestures

### 5.1 Post-Impact Motion Appears to Control the Illusion

[29] Much as analysis of a composition can lend new insight into its structure, analysis of the motions used by performers can be informative about the ways in which audiences experience this event. To determine the component of the gestures playing the strongest role in the illusion, participants performed the same duration rating task on three kinds of long and short gesture videos (two of which are shown in Example 9); pre-impact (showing only the motion up until the moment of impact, at which point they displayed a still image concurrently with the sound), post-impact (showing only the motion

concurrent with the sound), and full-gesture (i.e., the original full gestures). The magnitude of the illusion when viewing the post-impact gestures was similar to the magnitude when viewing the full-gesture videos—however the illusion was not found under the pre-impact condition (Schutz and Kubovy 2009b).

## 5.2 Moving Forward: Abstract Representations Using Point-Light Displays

[30] It is worth noting that the half-gesture videos in which the motion ceases upon impact could be said to lack a certain “ecological validity”<sup>(7)</sup>—they were created by artificially manipulating videos. Because it is not possible to view half-gestures in actual performances, it is not clear whether conclusions based on such impossible motions actually apply to musical situations. Therefore in order to ensure that our understanding of these gestures applies to actual music making, it was necessary to develop tools offering the ability to artificially render/edit the gestures without destroying their ecological validity.

[31] The first step in generating artificial-but-sufficiently-realistic gestures entailed creating “point-light” versions of the motions, which are demonstrated below.<sup>(8)</sup> Point-light animations are known to be effective representations of complex movements (Johansson 1973). Consequently, they are currently used to represent motions ranging from walking (Saygin et al. 2008), to drumming (Petrini et al. 2009; Arrighi et al. 2006), and dancing (Toiviainen et al. 2010).

[32] Our customized software offers the ability to render compelling point-light representations of the original long and short striking gestures using a flexible number of “joints” (Example 10). Experimental testing demonstrates that both the four-point and the single-point representations capture the salient aspects of the original gestures, yielding illusions that are statistically indistinguishable in magnitude from those found with the original videos (Schutz and Kubovy 2009a). Therefore these point-light representations of the long and short striking gestures are useful tools for creating realistic motion paths that can be rigorously manipulated.

## 5.3 How Do the Long and Short Gestures Differ in Their Motion Paths?

[33] Although it is clear that the long gestures “look long” and the short gestures “look short,” it is less clear which specific aspects of the physical movement of each gesture contribute to their influence on the perception of note duration. To explore the physical structure of these two gestures, we traced the vertical position of the mallet head (striking implement) over time, and plotted these values against one another. This previously unpublished data is shown in the left panel of Example 11. From this plot, it is clear that the gestures differ primarily in their motion post-impact, which is consistent with the earlier experiment demonstrating that videos containing only post-impact motion captured most of the illusion, whereas those exhibiting only pre-impact motion did not (Schutz and Kubovy 2009b; Section 5.1 of this paper).

## 5.4 The Use of Composite Gestures to Test the Primacy of Post-Impact Motion

[34] Point-light representations facilitate a number of experiments that would not be possible using only pre-recorded videos. One such experiment involved “hybrid gestures” mixing pre-impact motion of one gesture type with post-impact motion of another, permitting a direct comparison of their contributions to the illusion in a relatively ecologically valid context (Example 12). This experiment again found that post-impact motion was the principal driver of the illusion (Armontrout et al. 2009), building a stronger case that subsequent analysis of the gestures should focus on this segment of the motion. Much as analysis of harmonic progressions can illuminate mechanisms by which composers structure musical compositions to evoke certain listening experiences, analysis of these gestures can illustrate the ways in which marimbists structure their movements so as to shape their audience’s listening experience.

## 5.5 Duration of Movement is the Most Significant Component of Post-Impact Motion.

[35] The post-impact motion is complex, with the long and short gestures differing along a number of parameters—namely the distance covered by the motion and the time of the motion itself, not to mention its velocity, acceleration, and jerk (i.e., change in acceleration over time). These variables are intertwined in real-world motion, and picking apart their relative influence would be extremely difficult without the ability to control each parameter independently. Fortunately, single-dot

point-light displays offer such fine-grained control, affording opportunities to determine the relative contribution of each parameter (see **Example 13** for one of the three pairs of videos used in this experiment). Through a series of comparisons, Armontrout et al. (2009) determined that it is principally the duration of the post-impact motion—rather than its velocity, acceleration, jerk, or distance covered—driving this illusion.

[36] This finding constrains the number of variables requiring attention in a full analysis of these ancillary gestures—at least, with respect to understanding their effect on the perception of note duration. It also pinpoints the specific cues performers should focus on manipulating in order to apply these findings to music performance. Similarly, this specificity is useful to analysts interested in further explorations of the gestures themselves. Finally, these results neatly complete the previously described series of experiments aimed at deconstructing the complex motions used by one expert performer to solve an otherwise intractable musical problem—the marimba’s inability to control acoustic note duration.

[37] Together, these experiments demonstrate the challenges of fully understanding music listening in live performance situations (or when listening to audio-visual recordings) from musical notation or audio recordings alone. Additionally, they illustrate the benefits of a more holistic analysis involving rigorous consideration of the kinds of gestures used by expert performers to convey musical effects.

## 6. Implications and Interpretation

### 6.1 Summary

[38] The first section of this paper discussed the utility of ancillary gestures, noting that they play an important role in shaping the listening experience despite lacking acoustic consequences. The second focused on one specific type of ancillary gesture used to overcome acoustic limitations of the marimba; a “musical illusion” illustrating the distinction between acoustic information (i.e., “sound in the world”), and the way that sound is perceived in the mind of the listener/viewer. In order to put this phenomenon in context, the third section reviewed other illusions such as the McGurk effect—a compelling example of seeing influencing hearing. Both of these illusions illustrate that our brains often “think for us,” forming a coherent internal experience from a chaotic stream of sights and sounds.

[39] One cue used to trigger this seamless integration is causality—discussed in the fourth section. Manipulations weakening or destroying the causal cross-modal link between sight and sound correspondingly weaken or destroy the illusion itself. Finally the fifth section introduced point-light displays—a powerful tool for deconstructing and analyzing ancillary gestures. These displays not only capture the salient properties of the original gestures, but also afford the generation of life-like movements. Our customized software for rendering these motions allows for independent manipulation of all aspects of the striking gestures.<sup>(9)</sup> Therefore we have isolated and identified the specific property driving this illusion; the duration of movement in the post-impact motion.

[40] Collectively, these five sections review a series of experiments demonstrating that not only can ancillary gestures play a crucial role in the musical experience; they can also be analyzed, deconstructed, and interpreted in a rigorous manner. Therefore we believe that this work not only offers useful insight into music making, but also illustrates ways in which experimental research on music perception can supplement and augment traditional music-analytical approaches.

### 6.2 Significance for Music Theorists

[41] Much as this research has clear practical applications for performers and educators, we believe it can also inform analytical explorations of music’s formal structure.<sup>(10)</sup> For although certain properties of music can surely be studied either by examining the acoustic signal (i.e., the “sound”) or abstract representations of this signal (i.e., a musical score), our perception of these properties is affected by the implicit assumptions and interpretations discussed in [Section 3](#). Therefore, understanding the ways in which these assumptions, interpretations, and biases shape our perception of music is useful in understanding the structure of music itself.

#### 6.2.1 Implications for Analyses Related to Note Duration



[42] The fact that the durations notated on a score and realized in an acoustic signal might in some cases be perceived significantly differently in the minds of watching listeners holds significant implications for analyses involving note duration. For example, Selleck (1975) considered the relationship between duration and pitch in the Lutoslawski String Quartet. He discusses textural elements of the piece as being of primary interest, but also notes links between durational aspects and textural features. The gestures used to perform these notes are to some degree left to the discretion of the performer. Therefore, it is possible that gestures used to produce the sound (and consequently give rise to the duration-texture relationships) are shaping an audience's experience of this music in ways not immediately discernable from the score or acoustic signal alone. Although this composition is performed by strings rather than percussion, there is precedent for visual information playing a role in the perception of music performed on violin and viola (Gillespie 1997).

[43] Other analyses exploring note duration do in fact focus on music involving percussion—such as Boulez' *Le Marteau sans maître*. In an analysis of this composition, Winick (1986) noted that pitches of an ascending chromatic scale were generally associated with durations that increased incrementally. However, if the durations of the notes within the piece are *perceived* differently than notated because of performers' gestures, the relationships might in fact be different than suggested by the score alone. Given that keyboard percussion instruments play a significant role in this composition, percussionists have ample opportunity to employ ancillary gestures to great effect in sculpting their audience's listening experience.

[44] In another vein, Margulis (2007) explored a different aspect of duration in music—the duration of silence. Among other qualities, her study looks at the perceived duration of silences in a musical excerpt, including those used both before and after tonal closure. Her work illustrates that a silence's duration is indirectly affected by the perceived duration of the preceding event. If ancillary gestures are capable of altering perception of a note, this would consequently also potentially alter the duration of the succeeding silence.

[45] Although these examples explore the role of note duration in very different contexts and through different techniques (i.e., Selleck and Winick were basing their analysis on scores whereas Margulis explored a listener's perceptual experience), each of their outcomes and conclusions could be affected by ancillary gestures. Therefore, we believe that although scores may remain a starting point in music-analytical research, the experiments summarized in this review illustrate that the gestures used to perform written scores are musically relevant (and can make for interesting research topics themselves).

### 6.2.2 Implications for a Holistic Framework for Analysis

[46] In addition to the relevance for specific explorations focused on note duration, we believe experimental research on music perception holds general relevance for music analysis. In a sense, the value of exploring the relationship between impact gestures and perceived note duration is somewhat analogous to the value of using body movement to understand the perception of meter. In both cases the analysis lends insight into a psychological experience that is not otherwise directly observable.

[47] For example, metric structure is created through the experience of perceiving and anticipating rhythmic patterns (London 2004, 4), to the point that Gjerdingen (1989) has observed that “meter [is] a mode of attending.” Consequently, it is difficult to directly explore a listener's perceptual experience of metric structure without experimental studies, as the listening experience cannot be fully discerned from the acoustic signal alone. Therefore, techniques such as “music-induced motion” (Toivainen et al. 2010) discussed in Section 1.3 provide valuable insight, as they offer glimpses into the ways in which listeners organize music's structure. Analyses of the perceptual consequences of ancillary gestures employed by marimbists are useful for analogous reasons—they offer insight into the ways in which listeners perceive performances, information that is in turn useful for analysis and interpretation.

[48] Therefore, exploring the perceptual consequences of ancillary gestures allows analysts to better understand an audience's perception of musical notes. Although it is debatable as to whether the perceptual manipulations of duration explored in this research affect the theoretical “structure of the music” as opposed to the audience's “hearing of a particular performance”, this is a debate worth having. Moreover, it is a debate informed by experimental research on music perception and cognition.

Therefore we believe that recognition of music's multi-modal nature in fact holds potential for suggesting new directions and approaches for musical analysis.

### 6.3 Future Work

[49] Although the perceptual basis and physical correlates of this particular musical illusion are now understood, much work remains in order to fully explore its implications. To this end, we plan on recording professional marimbists performing musical excerpts using different kinds of gestures, then exploring these gestures' influence on music listening. Additionally, we are planning a complementary series of experiments exploring the use of ancillary gestures on other percussion instruments. Together, these studies will shed light on individual differences in percussionists' abilities to use gestures effectively, illuminating their musical utility.

[50] We are also interested in other ways in which extra-acoustic factors play a role in the musical experience, and are starting a new project exploring the effect of "moving to the beat" on listening to rhythmic music. Our preliminary results suggest that moving along while listening can both help improve a "listener's" understanding of music's rhythmic structure, as well as boost their confidence in these judgments—and likely their enjoyment of this information (Manning and Schutz 2011). Together, these lines of research will help to explore and document factors playing a role in the perception of music beyond those notated in a score. Ultimately, we hope this work will help inform our ability to play, analyze, and understand the musical experience. <sup>(11)</sup>

### Acknowledgements

This work was supported in part by funding from Natural Sciences and Engineering Council of Canada (RGPIN/386603-2010), the Ontario Early Researcher Award (ER10-07-195), and the McMaster Arts Research Board. We would like to thank Zachary Cairns, Peter Martens, Lee Hinkle, and two anonymous reviewers for helpful feedback/suggestions, and Brad Downie for assistance in preparing the video examples.

---

**Michael Schutz**  
McMaster University  
School of the Arts  
1280 Main Street West  
Hamilton, Ontario L8S 4M2  
[schutz@mcmaster.ca](mailto:schutz@mcmaster.ca)

**Fiona Manning**  
McMaster University  
School of the Arts  
1280 Main Street West  
Hamilton, Ontario L8S 4M2  
[manninf@mcmaster.ca](mailto:manninf@mcmaster.ca)

---

### Works Cited

- Adams, Wendy J., Erich W. Graf, and Marc O. Ernst. 2004. "Experience can Change the 'Light-From-Above' Prior." *Nature Neuroscience* 7, no. 10: 1057–58.
- Armontrout, James A., Michael Schutz, and Michael Kubovy. 2009. "Visual Determinants of a Cross-Modal Illusion." *Attention, Perception, and Psychophysics* 71, no. 7: 1618–27.

- Arrighi, Roberto, David Alais, and David Burr. 2006. "Perceptual Synchrony of Audiovisual Streams for Natural and Artificial Motion Sequences." *Journal of Vision* 6, no. 3: 260–68.
- Bailey, Elden. 1963. *Mental and Manual Calisthenics for the Mallet Player*. New York: Adler.
- Berry, Michael. 2009. "The Importance of Bodily Gesture in Sofia Gubaidulina's Music for Low Strings." *Music Theory Online* 15, no. 5. <http://www.mtosmt.org/issues/mto.09.15.5/mto.09.15.5.berry.html> (accessed 27 March 2012).
- Broughton, Mary, and Catherine Stevens. 2009. "Music, Movement and a Marimba: An Investigation of the Role of Movement and Gesture in Communicating Musical Expression to an Audience." *Psychology of Music* 37, no. 2: 137–53.
- Cage, John. [1940] 1976. *Living Room Music: Percussion and Speech Quartet*. New York: Henmar Press.
- Dahl, Sofia, and Anders Friberg. 2007. "Visual Perception of Expressiveness in Musicians' Body Movements." *Music Perception* 24, no. 5: 433–54.
- Davidson, Jane W. 1993. "Visual Perception of Performance Manner in the Movements of Solo Musicians." *Psychology of Music* 21, no. 2: 103–13.
- Floyd-Pautler, N. 2012. "Study Opens Doors for New Hearing and Autism Research." *The Hearing Lab*. <http://www.hearinglab.org/news/schutz.html> (published 12 February 2012, accessed 27 March 2012).
- Gillespie Robert. 1997. "Ratings of Violin and Viola Vibrato Performance in Audio-only and Audiovisual Presentations." *Journal of Research in Music Education* 45, no. 2: 212–20.
- Gjerdingen, Robert O. 1989. "Meter as a Mode of Attending: A Network Simulation of Attentional Rhythmicity in Music." *Integral* 3: 67–92.
- Hagen, Edward H., and Gregory Bryant. 2003. "Music and Dance as a Coalition Signaling System." *Human Nature* 14, no. 1: 21–51.
- Hatten, Robert S. 2004. *Interpreting Musical Gestures, Topics, and Tropes: Mozart, Beethoven, Schubert*. Bloomington, IN: Indiana University Press.
- Johansson, Gunnar. 1973. "Visual Perception of Biological Motion and a Model for its Analysis." *Perception and Psychophysics* 14, no. 2: 201–11.
- Kim, Jae Woo, Hesham Fouad, John L. Sibert, and James K. Hahn. 2009. "Perceptually Motivated Automatic Dance Motion Generation for Music." *Computer Animation and Virtual Worlds* 20, nos. 2–3: 375–84.
- Kleffner, Dorothy A., and Vilayanur S. Ramachandran. 1992. "On the Perception of Shape From Shading." *Perception and Psychophysics* 52, no. 1: 18–36.
- Krumhansl, Carol L., and Diana Lynn Schenck. 1997. "Can Dance Reflect the Structural and Expressive Qualities of Music? A Perceptual Experiment on Balanchine's Choreography of Mozart's Divertimento No. 15." *Musicae Scientiae* 1, no. 1: 63–85.
- Kubovy, Michael, and Michael Schutz. 2010. "Causality and Objects in Vision and Audition." *Review of Philosophy and Psychology* 1, no. 1: 41–61.
- London, Justin. 2004. *Hearing in Time: Psychological Aspects of Musical Meter*. New York: Oxford University Press.
- Manning, Fiona, and Michael Schutz. 2011. "Moving to the Beat Improves Timekeeping in a Rhythm Perception Task." In *Canadian Acoustics*, vol. 39, ed. R. Ramakrishnan and C. Laroche, 124–25. Quebec City: The Canadian Acoustical

Association.

- Margulis, Elizabeth Hellmuth. 2007. "Silences in Music are Musical Not Silent: An Exploratory Study of Context Effects on the Experience of Musical Pauses." *Music Perception* 24, no. 5: 485–506.
- McGurk, Harry, and John MacDonald. 1976. "Hearing Lips and Seeing Voices." *Nature* 264: 746–48.
- Petrini, Karin, Melanie Russell, and Frank Pollick. 2009. "When Knowing can Replace Seeing in Audiovisual Integration of Actions." *Cognition* 110, no. 3: 432–39.
- Ramachandran, Vilayanur S., and Diane Rogers-Ramachandran. 2008. "Touching Illusions." *Scientific American* 18: 60–63.
- Rosenblum, Lawrence D., and Carol A. Fowler. 1991. "Audiovisual Investigation of the Loudness-Effort Effect for Speech and Nonspeech Events." *Journal of Experimental Psychology: Human Perception and Performance* 17, no. 4: 976–85.
- Saldaña, Helena M., and Lawrence D. Rosenblum. 1993. "Visual Influences on Auditory Pluck and Bow Judgments." *Perception and Psychophysics* 54, no. 3: 406–16.
- Saygin, Ayse Pinar, Jon Driver, and Virginia R. de Sa. 2008. "In the Footsteps of Biological Motion and Multisensory Perception: Judgments of Audiovisual Temporal Relations are Enhanced for Upright Walkers." *Psychological Science* 19, no. 5: 469–75.
- Schutz, Michael. 2008. "Seeing Music? What Musicians Need to Know about Vision." *Empirical Musicology Review* 3, no. 3: 83–108.
- . 2009. "The Mind of the Listener: Acoustics, Perception and the Musical Experience." *Percussive Notes* 47: 22–28.
- Schutz, Michael, and Michael Kubovy. 2009a. "Causality and Cross-Modal Integration." *Journal of Experimental Psychology: Human Perception and Performance* 35, no. 6: 1791–1810.
- . 2009b. "Deconstructing a Musical Illusion: Point-Light Representations Capture Salient Properties of Impact Motions." *Canadian Acoustics* 37: 23–28.
- Schutz, Michael, and Scott Lipscomb. 2007. "Hearing Gestures, Seeing Music: Vision Influences Perceived Tone Duration." *Perception* 36, no. 6: 888–97.
- Selleck, John. 1975. "Pitch and Duration as Textural Elements in Lutoslawski's String Quartet." *Perspectives of New Music* 13, no. 2: 150–161.
- Stasack, Jennifer. 1987. *Six Elegies Dancing*. Everett, PA: Honey Rock Publications.
- Tan, Siu.-Lan, Peter Pfordresher, and Rom Harré. 2010. *Psychology of Music: From Sound to Significance*. Psychology Press.
- Thompson, William Forde, Phil Graham, and Frank A. Russo. 2005. "Seeing Music Performance: Visual Influences on Perception and Experience." *Semiotica* 156, nos. 1–4: 203–27.
- Toiviainen, Petri, Geoff Luck, and Marc R. Thompson. 2010. "Embodied Meter: Hierarchical Eigenmodes in Music-Induced Movement." *Music Perception* 28, no. 1: 59–70.
- Van Camp, Julie Charlotte 1981. "Philosophical Problems of Dance Criticism." PhD diss., Temple University. <http://www.csulb.edu/~jvancamp/diss.html> (accessed 27 March 2012).
- Vines, Bradley W., Carol L. Krumhansl, Marcelo M. Wanderley, and Daniel J. Levitin. 2006. "Cross-Modal Interactions in the

Perception of Musical Performance.” *Cognition* 101, no. 1: 80–113.

Wanderley, Marcelo M. 2002. “Quantitative Analysis of Non-obvious Performer Gestures.” In *Gesture and Sign Language in Human-Computer Interaction*, ed. I. Wachsmuth and T. Sowa, 241–53. Berlin: Springer Verlag.

Wanderley, Marcelo M., Bradley W. Vines, Neil Middleton, Cory McKay, and Wesley Hatch. 2005. “The Musical Significance of Clarinetists’ Ancillary Gestures: An Exploration of the Field.” *Journal of New Music Research* 34, no. 1: 97–113.

Winick, Steven. D. 1986. “Symmetry and Pitch-Duration Associations in Boulez’ *Le Marteau sans maître*.” *Perspectives of New Music* 24, no. 2: 280–321.

---

## Footnotes

1. *Critical band* is a psychoacoustic term describing the pitch distance within which two tones “conflict.” For humans, critical band width is wider in the lower frequency ranges, and root position triads in lower registers therefore sound “muddy.” Although this phenomenon is often attributed to differences in the register (i.e., pitch height) of the triads, register is really the *proximate*, rather than the *ultimate* cause.

The distinction between these types of explanations can be seen most clearly through a parallel example. For instance, the *proximate* explanation for our desire to consume desserts is that the instant gratification of sweet foods is highly pleasurable. However, this account really describes, rather than explains the phenomenon of interest. The *ultimate* explanation, rather, is that in our ancestral environment, sweet foods were both rare and nutritionally valuable—therefore our predecessors evolved an intense affinity that has passed forward through generations. Similarly, the register difference between the high and low triads is the proximate cause of the differences in their perception, whereas the ultimate cause is that the differential sizes of critical bands in low vs. high registers result in different perceptual experiences when listening to the triads—even when they share an identical intervallic structure.

[Return to text](#)

2. That is not to say the use of theatrical gestures is confined to percussion alone. Sofia Gubaidulina’s output includes multiple works for cello and bass, as well as extended sections of compositions consisting entirely of ancillary gestures—i.e., without any intended acoustic effects (Berry 2009).

[Return to text](#)

3. In order to focus primarily on the implications of this work, technical details about experimental design and statistical analysis are kept to a minimum. Readers interested such details can find them in Schutz and Lipscomb 2007.

[Return to text](#)

4. Research on the well-known “light from above” illusion is summarized in an article in *Scientific American* titled “Shading Illusions: How 2-D Becomes 3-D in the Mind” (Ramachandran and Rogers-Ramachandran 2008). This article is available online at <http://www.scientificamerican.com/article.cfm?id=seeing-is-believing-aug-08>.

[Return to text](#)

5. This widely-used demo is also available on YouTube (<http://www.youtube.com/watch?v=aFPtc8BVdjk>).

[Return to text](#)

6. Each experiment in this paper used a new group of participants.

[Return to text](#)

7. Ecological validity refers to the degree to which an experiment accurately mimics real world situations. The use of videos depicting human motion creating acoustically natural sounds means these studies are quite “realistic” relative to the bulk of research on psychophysics. However, we recognize that half-gestures are clearly somewhat “unrealistic” from a musical

perspective.

[Return to text](#)

8. Traditional point-light displays involve only points of light, rather than a skeleton connecting the dots as in the virtual marimbist used in these experiments.

[Return to text](#)

9. This software is part of a complex suite of tools we are developing to facilitate experimental research on perception and cognition. Once fully developed we will share these tools with other interested researchers through our website at [www.maplelab.net/software](http://www.maplelab.net/software).

[Return to text](#)

10. Additionally, we are currently exploring clinical applications of this work through collaborations with clinical psychologists/audiologists. For example, we are working with autism expert Dr. Laura Silverman at the University of Rochester on an NIH funded project using this paradigm to explore sensory integration dysfunction in individuals with autism spectrum disorder. For more information see *Study Opens Doors for New Hearing and Autism Research* by Nedra Floyd-Pautler on the Hearing Lab website at <http://www.hearinglab.org/news/schutz.html>.

[Return to text](#)

11. For information on this and other lab projects (including samples of videos for use in class demonstrations) visit the MAPLE (Music, Acoustics, Perception and LEarning) Lab online at [www.maplelab.net](http://www.maplelab.net).

[Return to text](#)

---

## Copyright Statement

**Copyright © 2012 by the Society for Music Theory. All rights reserved.**

[1] Copyrights for individual items published in *Music Theory Online (MTO)* are held by their authors. Items appearing in *MTO* may be saved and stored in electronic or paper form, and may be shared among individuals for purposes of scholarly research or discussion, but may *not* be republished in any form, electronic or print, without prior, written permission from the author(s), and advance notification of the editors of *MTO*.

[2] Any redistributed form of items published in *MTO* must include the following information in a form appropriate to the medium in which the items are to appear:

This item appeared in *Music Theory Online* in [VOLUME #, ISSUE #] on [DAY/MONTH/YEAR]. It was authored by [FULL NAME, EMAIL ADDRESS], with whose written permission it is reprinted here.

[3] Libraries may archive issues of *MTO* in electronic or paper form for public access so long as each issue is stored in its entirety, and no access fee is charged. Exceptions to these requirements must be approved in writing by the editors of *MTO*, who will act in accordance with the decisions of the Society for Music Theory.

This document and all portions thereof are protected by U.S. and international copyright laws. Material contained herein may be copied and/or distributed for research purposes only.

*Prepared by John Reef, Editorial Assistant*