ANNABEL J. COHEN

Film Music
Perspectives from Cognitive Psychology

✸

In the language of cognitive psychology, cinema is a multisensory stimulus that, millisecond by millisecond, impinges on sensory receptors and excites networks of neuronal activity in the mind of every film spectator. Writing in 1936, long before the emergence of a rigorous cognitive psychology, literary critic Walter Benjamin (1969) described the daunting perceptual-cognitive challenge that cinema presents to each member of the audience:

> Let us compare the screen on which a film unfolds with the canvas of a painting. The painting invites the spectator to contemplation; before it the spectator can abandon himself to his associations. Before the movie frame he cannot do so. No sooner has his eye grasped a scene than it is already changed. It cannot be arrested. . . . The spectator's process of association in view of these images is indeed interrupted by their constant, sudden change. (240)

With its reference to the visual domain, Benjamin's comparison of the vicissitudes of the cinema screen to the painter's static canvas is apt as far as it goes. But no account of cinematic experience is complete without consideration of the auditory domain, that of music, speech, and sound effects. For example, Michel Chion emphasizes that what we see is always altered by what we hear. He refers to the "added value" of sound that "engages the very structuring of visions . . . by rigorously framing it" (Chion 1990, 7). Cognitive psychology enables us to examine in an experimentally controlled way how joint stimulation from the auditory and visual domains operate in our cognition of the cinema.

To illustrate the use of, and perhaps need for, a cognitive approach to film music, consider a brief sequence from *Parenthood*. While depicting an ordinary Little League baseball game, the camera tracks the trajectory of a baseball and focuses finally on a young player. This is his moment. Will he catch the ball or not? First he fumbles and then miraculously recovers the ball. On the surface, the excerpt depicts an everyday childhood baseball game, but Randy Newman's music tells us that this is not an ordinary event. Rousing triplets in the full orchestra, along with unresolved harmonies, accompany the ball's flight. As the boy catches the ball, the music ceases; the umpire yells, "He's outa there," immediately followed by a jubilant brass fanfare in $\frac{4}{4}$ time. Can there be disagreement that these contrasting music selections produce associations that add meaning to the film? Surely not. The rousing triplets and unresolved harmonies heighten the audience's understanding of the emotional significance, but also anxiety and uncertainty, of the young boy and his father. The fanfare celebrates the boy's success, and so too then does the audience. In this example, film music clearly adds meaning for the audience.

Suppose we now play back only the musictrack from the excerpt. The music might conjure up images (past visual associations) of the ball, the boy, the fumble, or the catch. Thus, a second role of music is to aid the audience in remembering parts of the film, reinforcing the connection among film events and eliciting associations.

Without music, the images seem prosaic, mundane, even lifeless; with music, however, the world of film comes alive. Through music however, film spectators become part of the crowd watching the game, forgetting the theater seat and the screen. It is ironic that music is such a powerful vehicle in this regard. For most of the excerpt under discussion, diegetic sound effects are absent: only music is heard behind a "silent" crowd. In the real-life situation, the opposite would obtain: ambient crowd noise and no music in the background. In the film, as soon as the boy catches the ball, diegetic sound effects return. A psychological sense of reality continues in spite of the discontinuity in the audio background. Does the film audience miss the diegetic sound? Does the audience even notice the absence? Does the music make the event more compelling?

This examination of one fairly representative excerpt from a Hollywood film points to three functions of film music: (1) music interprets and adds meaning; (2) it aids memory; (3) it suspends disbelief. But it is one thing to postulate these functions of film music speculatively and another to subject the postulates to convincing empirical tests. Here is where the methods of cognitive psychology prove useful. With its reliance on controlled experiments, cognitive psychology can place a provocative but speculative theory on solid psychological foundations. Empirical evidence from psychology

In keeping with the standards of psychology literature, social-sciences citation style is used in this chapter. *Eds.*

experiments helps complete our understanding of how music aids the audience's comprehension of a film.

Meaning

It is generally assumed that musical soundtracks influence the interpretation of images they accompany. For example, Royal Brown (1994) describes music's tendency to narrativize film images. He makes the general assumption that "musical scores very often tell our emotions how to read a given filmic sequence" (54). Similarly, George Burt (1994) writes of music's power to "open the frame of reference to a story and to reveal its inner life in a way that could not have been as fully articulated in any other way" (4). Showing how a particular passage of music influences an interpretation of a film clarifies the statement that "film music influences interpretation." Possible contributions of cognitive psychology to this issue are reviewed below.

For instance, one of my experiments investigated a single dimension of emotional meaning (happy/sad) as a product of the interaction between a monophonic melody and one moving object. The musical examples used repeating broken major triads that differed in only two respects: tempo and pitch. These musical examples were presented at a slow, moderate, or fast tempo and were low, medium, or high in pitch. Previous work in the auditory dimension alone had revealed that tempo and pitch height controlled listeners' judgments on a five-point happiness/sadness rating scale (Trehub, Cohen, and Guerriero 1985). In other words, low pitch on average led to a rating below three and high pitch led to a rating above three. Slow and fast tempo led to low and high numerical judgments, respectively. The visual stimulus in the audiovisual experiment was a computer-generated ball that bounced at one of three heights and at one of three speeds. Again, perhaps surprisingly, previous work had suggested that these dimensions led to systematic judgments on the happy-sad dimension. Low, slow bounces were judged as sad, and high, fast bounces were judged as happy (Cohen 1993).

As might be expected, the experiments show that when the auditory and visual dimensions are congruent (low bounce/low pitch, or high bounce/high pitch), the judgment is consistent with the presentations of either audio or visual modality alone. When the audio and visual dimensions diverge, however, the judgment tended to fall between the rating for either the audio or visual dimension alone. A slow melody in a low register, for instance, produced judgments of the high-fast ball that were lower (less happy) than they were in the absence of the music, even when subjects were asked to judge only the ball. They seemed unable to resist the systematic influence of music in determining their interpretation of the image. This suggests that, in the simplest case, the cognition of film music is additive: it sums up associations or meanings mentally generated by the different film and music components.

Using slightly more complex video and audio materials, another study measured multiple scales of affective meaning employing the semantic differential technique (Marshall and Cohen 1988).[1] Here, the video material was an animation involving three geometric objects, a small circle and a small and a large triangle. In general, viewers tended to interpret the action of the inanimate objects in a stereotypic and anthropomorphic way; for example, typically viewers see a large bully persecuting a loving, innocent couple. The experiment was designed to test whether and to what extent music alters this basic meaning of the animation. Two contrasting musical scores, A and B, were judged, using rating scales, by different groups of listeners, and a third group judged the film.[2] Then, with either music A or music B in the background, two other groups judged the film overall and each of the three geometric characters. The ratings of the film overall differed with respect to the soundtrack, as might be expected; of special interest, however, was that the apparent effect of music extended to judgments of the geometric objects. The small triangle, for example, seemed more active with one of the musical scores. It is likely that congruence between the temporal patterns of the latter music and the motion of the geometric figures drew attention to the small triangle, which permitted musical meanings to become linked with the small triangle. If this hypothesis, which forms the basis of the "Congruence-Associationist framework" discussed below, is correct, then different music could elicit a different visual focus of attention.

Studies using live-action films suggest that the Congruence-Associationist hypothesis extends to a more typical filmgoing experience. In one experiment (Cohen 1993), two video excerpts—in one, a woman runs from a man, in the other, two men fight—were selected from a film by an amateur director, and two contrasting music selections were used. Once again the film and two music selections were evaluated by independent groups of viewers. Here, the results were mixed: although music strongly affected the interpretation of the male-female chase, the music for the fight scene did not have a strong effect on viewer interpretation. In retrospect, it is clear that the extent of musical influence depends to a large extent on the degree of visual ambiguity. The fight scene was more highly determined than the male-female interaction, which could be interpreted as representing either an aggressive interaction or amorous play. Here ambiguity was probably

created by the cultural understanding of gender roles; presumably, music resolves some of the ambiguity of the situation. A final experiment in this series confirmed that music helps determine meaning in an ambiguous situation (Bolivar, Cohen, and Fentress 1994). Here, the visual examples consisted of wolf interactions in which it is unclear whether the wolves are playing or fighting. In this case, too, music was found to have a direct impact on judgments concerning the character of the interaction.

Thus, all four studies reviewed above (the bouncing ball, geometric animation, human interactions, and wolf interactions) found that music helps define the meaning of a scene and that the effect of music is most pronounced when the situation depicted is ambiguous. At the same time, statistical analysis suggests that associations of image are processed independently of musical associations, and that the final meaning we receive from a film is additive, the result of the total associations generated, a finding consistent with a modular theory of mind, which will be discussed below. Not only do these empirical findings of cognitive psychology assure us that film music influences film interpretation, but they also tell us something about what the various components of the music contribute, how much they contribute, and under what circumstances, such as visual or narrative ambiguity, this contribution is maximized.

Memory

A second common assumption of film-music scholarship that can be given empirical support using the method of cognitive psychology concerns effects of musical memory for film. For example, Claudia Gorbman (1987) states that, in just one presentation, a musical leitmotif comes to represent the protagonist Mildred Pierce, in the film of the same name (1945). Here, Gorbman makes an assumption about how musical memory works. Similarly, in Woody Allen's *Oedipus Wrecks* from the *New York Stories* trilogy, the famous ballad "All the Things You Are" is presented during two interactions between a middle-aged wimpy lawyer (played by Allen) and a loony psychic (Julie Kavner) who has taken more than a professional interest in the male protagonist. Does a third presentation of the song while Allen is alone (fondling a roasted chicken leg which Kavner had given him) elicit the audience's memories of Kavner? What evidence is there that music transfers its meaning to another similar situation, that music and film become integrated in the mind, or that the music brings back images of the film and acts as a retrieval cue for the film? Such results, which seem quite intuitive, are in fact somewhat difficult to test empirically.

Nevertheless, aspects of these claims can be empirically tested. One study (Cohen 1995) concludes that music does little to aid in the memory of visual cues because visual recognition is sufficiently high that supplementation is unnecessary. At the same time, the study also demonstrates that subjects are quite capable of remembering pairings of music and images. What then is the function of these music and image pairings? Boltz, Schulkind, and Kantra (1991) found that music aids in directing visual attention; music, in short, marks images for conscious attention, but the extent to which it does so depends on timing and congruence in meaning. In this study, the music either accompanied a scene's outcome and thereby accentuated its affective meaning, or foreshadowed the same scene and thereby created expectations about the future course of events. The background music was either congruent or incongruent with the episode's outcome. In a subsequent recall task, violations of expectation facilitated memory in the foreshadowing condition, while mood-congruent relations led to better performance in the accompanying condition. Results from a recognition task further revealed that music helped subjects recall scenes they did not otherwise remember. The findings were discussed in terms of the ability of background music to manipulate visual attention depending on whether the music either precedes or accompanies the visual image, and whether it is congruent or incongruent with the image. Such evidence, together with the positive results from explicit questioning (Cohen 1995), supports the conclusion that music may be linked with image information in a single presentation, a conclusion that the intuitive model of the cognition of leitmotifs presupposes.

Furthermore, the effects of music on memory for images may extend beyond those discussed above. For example, Thompson, Russo, and Sinclair (1994) show that music influences judgments about perceived closure in a film narrative. They note that in nonmusical contexts, closure has been found to influence memory, and they suggest that future studies of music on closure in film could be directed also to memory issues.

Suspension of Disbelief

Gorbman has stated that "music removes barriers to belief; it bonds spectator to spectacle, it envelops spectator and spectacle in a harmonious space" (1987, 55). From a cognitive standpoint, Gorbman's statement raise an interesting question: how could music produce this profound effect of suspending disbelief? If recall of visual material is neither worse nor better when accompanied by music; and if music is equally well retained whether or not it is accompanied by images, then it appears that musical information does

not significantly compete with visual information for mental-processing resources. Likewise, it has also been shown that music does not interfere with the cognitive ability to process dialogue or vice versa (Cohen, unpublished). It is possible therefore that music is processed independently of both film and speech at the early stages of perception, and only at later stages of cognition is the information integrated. If this is so, music requires mental resources beyond those needed to process either image or dialogue. One explanation for the presence of music in film might be that, to some extent, this increased activation heightens our sense of the diegetic film world as real. The real world typically stimulates auditory and visual senses simultaneously. Thus, mentally processing music and film creates a situation mentally more similar to that of real life than processing film alone. But music does not typically accompany real-life events as it does in film. So it may also be that the affective associations of music contribute something lacking in the reality constructed by the film.

As shown in behavioral research of Bolivar et al. (1994), subjects can process music and visual material simultaneously when asked to do so and they can also direct their attention to one or the other modality. Anecdotally, filmgoers are typically unconscious of most film music. Using a survey, Archie Levy (personal communication 1990, 1993) showed the extent to which music escapes notice in films. He asked people leaving a theater what they thought of the music. Most people thought it was fine. Not a very surprising comment, until it is noted that the film had no musical cues except under the opening credits.[3] While music in film serves as a vehicle used to transport emotional meaning, it is a vehicle that is often "inaudible" (Gorbman 1987), much as the font of this page is transparent until I draw attention to it. We can discriminate between `Courier` and *Galliard,* but when reading, we don't really much care whether it is one or the other, as long as it is legible. Similarly, the viewer-listener accepts the musical meaning, but acoustical properties of the music itself seem to function transparently as a kind of "acoustical font."[4]

The situation may be analogous to cognitive psychologist Anne Treisman's concept of illusory conjunction in visual information processing (Treisman and Schmidt 1982). Treisman rapidly presented visual elements on a computer screen, for instance, short lines or simple forms in various orientations and in various colors. Forms and colors were presented only in some colors or orientations. For example, a line tilted at sixty degrees might appear as blue but never as red while a line tilted at ninety degrees would be red but never blue. Subsequently, subjects were presented with visual objects and asked how sure they were that various combinations of features had been previously presented. Subjects often claimed

they had seen objects that in fact had not been presented. It is true that the parts of the objects may have been seen, for example, blue color, square shape, sixty-degree orientation, but the parts had never been presented together. That the parts were previously seen together, however, is an illusion.

Something similar may happen with music and images. In this analogy, we can consider music to have two components: an affective component and an acoustical, structural component. When these two components of music are presented simultaneously with a visual image, the conjunction of the affective element and the visual image makes a new meaningful whole, a whole much closer to our sense of reality than the visual image alone, or than the visual image conjoined with both affective and acoustical components. Through the illusory conjunction process, the affect, originally carried via the acoustic properties of music, attaches to the visual stimulus. One or both of two hypothetical processes may account for this disjunction of acoustical and affective information. First, a reality test may exist that must be passed in order for information to achieve conscious attention. The visual information plus the affect from music can pass this test because in real life visual information is often associated with such feelings. For example, when children are playing Little League baseball, spectators feel sad when the child fumbles the ball. (For unknown reasons, a visual depiction of such a scene does not generate this emotional response on its own.) Thus, the visual scene plus the feeling of sadness passes the reality test. However, visual information plus the acoustic aspect of music would not pass the reality test, because the combination of musical sounds and Little League baseball does not match reality. The match to reality is important because it may relate to psychological attentional processes. Many things vie for our attention, and Grossberg (1995) argues that only those that match typical expectations receive attention. He has proposed an Adaptive Resonance Theory (ART) of attention that entails such a mechanism whereby information in short-term memory proceeds to consciousness only if matched by predictions based on "realities" of long-term memory. Such predictions can be based on the schemas we have developed through experience and hold in long-term memory in order to interpret moment-to-moment experiences. The same idea, however, applies to arbitrarily constructed experimental stimuli (like geometric objects of different colors and orientations).

A second but not mutually exclusive mechanism for how music might aid film in constructing its reality arises through overloading the mental processes responsible for the separate visual and auditory senses. The flood of stimulation provided by music and film may somehow reduce the criteria for reality—anything creating this much mental activity must be taken seriously though perhaps in the same way as a dream suspends disbelief

(Mitroff and Bennis 1993, 56–58). An extension of this is the notion that the less realistic the images, the more music must contribute to a suspension of disbelief; for example, music is more likely to need to "suspend disbelief" in a cartoon than a documentary. These ideas are speculative, but admit to further testing.

Cognitive psychological studies of memory and attention can thus begin to account for the mind's ability in the cinematic context to exploit the semantic, affective dimension of music while ignoring the acoustical aspect not anchored in the image. Subjects in experiments can seem oblivious to the acoustical dimension of music (although they will focus on it and remember it when asked to do so, in the same way that we see the font when asked about it). A cognitive theory such as ART can account for the failure of the acoustical dimension of music to reach consciousness.

Modularity, Domain Specificity, Music, and Cinema

Film music encourages us to consider the extent to which certain domains of experience are addressed distinctively by the brain. As a clear example, it has long been known that a part of the temporal cortex called Wernicke's area, in the left hemisphere in right-handed individuals, specializes in the cognition of speech (reviewed in Patel and Peretz 1997; Peretz 1993). There is a class of sounds that, if interpreted as speech, are processed in the left hemisphere but, if interpreted as nonspeech music, are processed elsewhere. Sounds are apparently sent to Wernicke's area to be processed as speech and to somewhere else to be processed as music. Evidently, the brain treats certain information on the basis of its function (Dingwall 1993). Whereas this suggests separable and independent modes of mental processing, there is also evidence that information from these different sensory inputs can be integrated at higher levels of cognition. Several studies have shown that song lyrics, for example, provide memory retrieval cues for the music associated with them, and vice versa (Crowder 1993; Crowder, Serafine, and Repp 1990). Something similar likely occurs with film music, as it has been demonstrated that music can alter the interpretation of visually presented action (Cohen 1993) and even influence the extent to which a filmed action appears completed (Thompson et al. 1994).

The relative cognitive independence of music and image seems consistent with Fodor (1983), who argues that the sensory input systems must be modular. Modularity of the mind starts with a fairly obvious fact of physiology: the peripheral visual system encodes activities that take place at the retina, not at the cochlea. Modularity has a number of important consequences, one of which is that not all the information obtained from the perceptual and cognitive activity of peripheral modular processes can be recovered through conscious reflection. The outputs of the modular input systems likely connect in order to make a holistic experience. This means that the holistic experience is synthetic, that is, an illusion. Our perception of reality is a construct to the extent that it appears integrated. In the fifteen years since Fodor's celebrated treatise on modularity, a number of neurophysiological and psychological studies have provided additional evidence on the extent of separable mental processes. Some of these are reviewed below as being relevant to our concern about how the complexity of music can be handled amid the simultaneous processing of other complex media.

Historically, the first evidence of a special brain area for music was revealed by Milner (1962) in patients who had only one cerebral hemisphere intact. An undamaged right hemisphere, as compared to left, was associated with better performance on a Seashore task of musical memory. Similar results were obtained using a more extensive music test battery from Gordon (Kester et al. 1991). Studies of patients with unilateral temporal lobectomy have also revealed enhanced learning and retention of melodic material when the right hemisphere is preserved (Samson and Zatorre 1991). Other studies with patient populations have shown that melodic imagery and perception (Zatorre and Halpern 1993) and timbre sensitivity (Samson and Zatorre 1994) are lateralized to the right hemisphere. In normal listeners, lateralization of musical and verbal functions has been shown through dichotic listening tasks (different sounds to two ears). A left-ear (right hemisphere) advantage for musical information and a right-ear (left hemisphere) advantage for verbal information is typically reported (Kimura 1964, 1967). Although information from each ear goes to both hemispheres, the ear advantage implies privileged processing by the contralateral hemisphere. PET (Positron Emission Tomography) scan studies conducted by Zatorre and his colleagues further support the notion of special brain loci for music.

Results have revealed that the brain activity is systematically related to the type of musical task (Zatorre, Evans, and Meyer 1994). In other words, different parts of the brain "light up" when certain kinds of listening occur. The part of the brain involved in rhythmic activities is a more primitive part of the brain, the cerebellum, which is associated with movement and balance (Jourdain 1997). It is an area that would also be involved in dance and motor responses to the music. When a song is presented with lyrics, part of the right hemisphere is stimulated by the music and part of the left hemisphere is stimulated by the words. This result is fully consistent with the work of Milner and Kimura. Zatorre, Halpern, Perry, Meyer, and Evans (1996), which also showed that musical perceptual and imaging tasks activate almost identical areas in the temporal cortex and supplementary

motor area. In other words, musical (that is, auditory) images (in the absence of external stimulation) and music perception (auditory representations accompanying stimulation) entail the same brain activity.

The brain-imaging tasks confirm that several aspects of musical processing occupy different areas of the brain than does speech processing. This is consistent with the idea that processes considered to represent distinct media, such as speech and music, do not compete for the same brain space. Separate, independent systems for verbal and musical information are also implied in psychological studies of short-term memory by Deutsch (1970). In these studies, subjects on each trial heard one tone followed by a comparison tone and were asked whether the second tone was the same as the first tone. This task was not difficult, but, if a sequence of different tones was interpolated between the first and last tones, then the task became more difficult. The interpolated tones interfered with the memory for the first tone. However, when a series of spoken numbers replaced the interpolated tones, memory comparison of the first and last tones once again became easy. This demonstrates that tones were processed by the brain independently of verbal information, like spoken numbers.

The psychological evidence that different media, in this case speech and music, may excite independent brain activities is consistent with the views on intelligence of Gardner (1983/1993), who has argued for eight types of intelligence, music being one of these.[5] Other intelligences relevant to the present discussion are linguistic and spatial. Modularity and the notion of separate intelligences are consistent with the view that music activates separate parts of the mind distinct from visual-spatial and verbal information, the chief aspects of the cinematic experience.

The neurophysiological and psychological studies taken together support the view of domain-specific mental structures for storage and processing of visual, verbal, and musical information. This provides a foundation for understanding how music can function in a variety of ways in cinema.

Congruence-Associationist Framework for Understanding Film-Music Communication

A preliminary framework representing some aspects of the processing of cinema is presented in figure 1. The framework is essentially a flow diagram that highlights certain aspects of the independent and interactive processing of the three primary media in cinema. It also illustrates that, through structural congruence, music directs specific visual attention and conveys meaning or associations (Cohen 1990, 1993). Hence the framework is referred to as Congruence-Associationist.[6] It emphasizes that music is a vehicle transporting a variety of information, only some of which is relevant to a particular cinematic goal. The brain seems to be able to select what is useful for the goal at hand. A prime example is the role that affective meanings or associations from music provide to a visual narrative; in this case, the sounds themselves are almost of no concern whereas the affective meaning conveyed by the music is often of critical importance.

The diagram represents six levels of information flow, from A to F. Level A represents the surface structure—of speech, video, and music. At level B, each of these is analyzed into components by domain-specific systems. For music, this means decomposition into temporal structures and emotional meaning (affect). (There are, of course, other kinds of feature analysis that would take place, but the diagram emphasizes these two.) Level B affords the possibility for cross-modal congruencies to take control and automatically lead preliminary attention only to a portion of the visual information, shown here as the material within the oval.

To explain further: cognitive psychologists typically have applied Gestalt principles to visual pattern and lately to auditory information (for example, Bregman 1990, 1993; Narmour 1991). Rarely are the principles applied to the two domains at once. But film music provides an opportunity for this application. It is my view that music accompanying a film automatically elicits bottom-up principles that entail grouping across auditory and visual domains. When the auditory information and visual information are structurally congruent, the visually congruent information becomes the figure, the visual focus of attention. What I am suggesting is that Gestalt theoretic ideas (I have avoided the word "principles" here because it implies more precision than is actually the case), which are typically applied to visual or auditory domains independently, can be applied equally productively to conjoint visual and auditory dynamic information. It follows that, through innate Gestalt grouping processes, music can define the visual figure against the audiovisual background; music can sometimes determine what is the visual focus of attention.

These ideas relate to those in the film-music literature on sensitivity to, and effectiveness of, congruent musical and film structures. An example is Prokofiev's score for the famous "Battle on the Ice" sequence in *Alexander Nevsky,* where the temporal contour of the melody mimics the static visual contour of the scene. "Mickey-mousing" supplies countless other examples. In these instances, the composer or director has intuitions that similar formal or structural congruence between music and video patterns promote an appreciation for the content of the film. The present framework enables research on the following questions: To what extent is the cross-dimensional structural congruence encoded? How accurate must the congruence be for

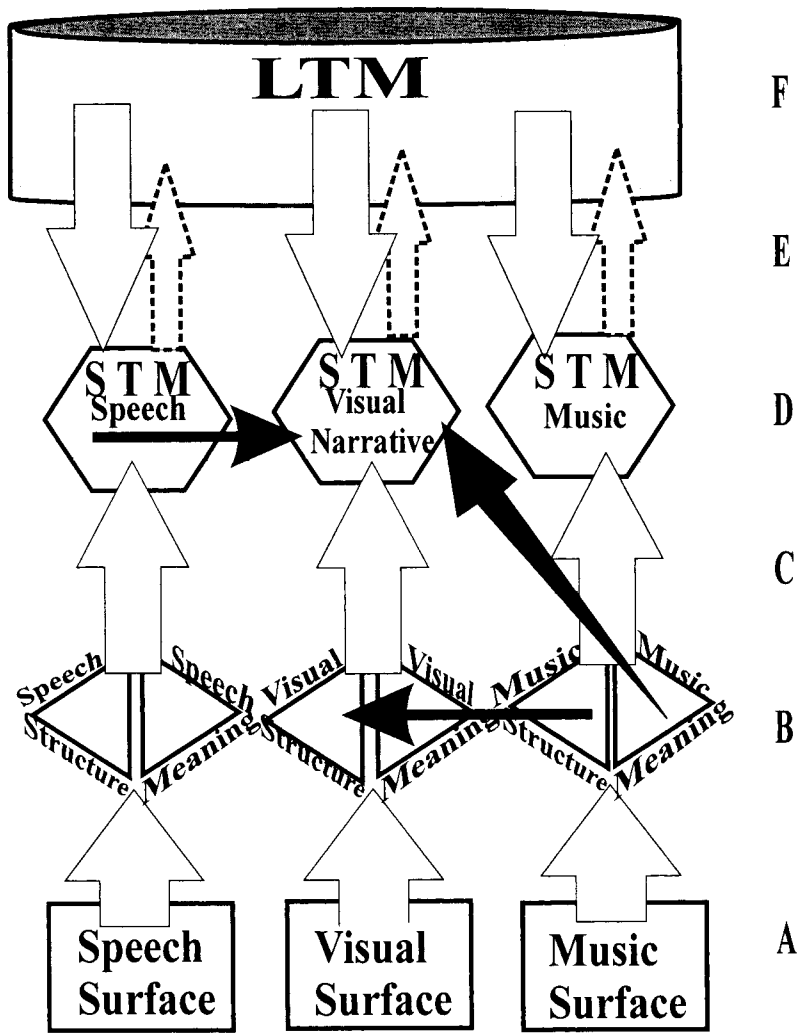

Figure 1. *Congruence-Associationist Framework* for understanding film-music communication. Information-flow diagram of mental activity underlying the joint processing of continuous media of speech, video, and music. At A, the physical surface features of each domain are represented. At B, the surface is analyzed into structural and meaning components. Musical structural information combines with similar visual structural information such that corresponding visual patterns receive priority in traveling (via C) to short-term memory (STM) at D. Level D represents modular STM systems for speech, visual narrative, and music. Note that information from music meaning (at B) travels both to the visual narrative STM at D and to the music STM at D. Information in STM is attended if matched by information sent from long-term memory (LTM) at F. Hypotheses to make sense of the emotional, speech, and visual information are generated at F to match the visual narrative STM at D. Some of this information at D may travel back to LTM (via E) to form new long-term memories at F. Attention to the music is also supported by the analysis through synthesis and a matching approach represented by levels D, E, and F. A similar process is envisioned for speech.

it to be effective? What dimensions in the auditory and visual domains can be manipulated to establish congruence? To what extent does audiovisual congruence focus visual attention?

Returning to the remainder of the framework of figure 1, information is transferred by level C to short-term memory (STM) at D. Certain visual information has priority of transfer, specifically, that in the gray oval. In other words, not all of the visual information that is potentially available reaches short-term memory. In addition, information on musical meaning is transferred not only to a music STM station but also to the visual STM station. The visual STM station is referred to as the STM Visual Narrative. The term *narrative* indicates the goal of the STM process to make sense of the visual information using whatever information is at hand. Thus, the affective quality of music is directed here because it is useful in determining the meaning of the visual scene.

To explain how we become conscious of this material in STM, we turn to Grossberg's (1995) model of conscious attention. This model asserts that material in STM must be matched by material in long-term memory (LTM). Assuming that experiences in LTM include affective tone, it is conceivable that a matching process would take place for visual and affective information from a film, whereas the acoustical properties of the musical accompaniment that gave rise to the musical affect would not be matched. In other words, material in STM would be matched by material generated by hypotheses from LTM at F. This would explain why the acoustical aspects of the music would not generally be attended to: the acoustical aspects of the music do not make sense to LTM (where is that background music coming from?), and no hypotheses would be generated easily to include it (unless, of course, music were part of the diegesis, for example, attending a concert or taking a music lesson). Thus, the main phenomenal experience (E) is one of a narrative with visual and emotional components. Running in parallel, all aspects of the music can be processed at a conscious level (see levels D, E, and F in the music column), as it is known that simultaneous tasks can co-occur (Neisser and Becklin 1975), and there is evidence that background music is remembered (Boltz et al. 1991; Cohen unpublished). (A similar process is envisioned for speech as well, but this is not the focus of the present essay).

This framework explains the puzzling and paradoxical role of background music in film. Music adds information that is both consistent and inconsistent with the narrative. The affective quality is consistent; the acoustical aspects of the music are not. Although the affective associations produced by the music seem to belong to the corresponding images, the sounds that produced those associations do not. Somehow, the brain

attends to this affective meaning, while ignoring or attenuating its acoustical source (see Cohen 1994). Fodor (1983) and many others have commented that only some of the information that is encoded reaches consciousness. In the case of background music, it seems that the affective meaning of the music reaches consciousness but the acoustical musical surface does not, at least in the stream focusing on the visual interpretation. Of course, there may be exceptions to this general notion. Some individuals may in fact be particularly sensitive to film music and actually attend films because of it, just as a minority of individuals have absolute pitch and are aware of the actual note name and frequency of musical tones they hear.

Conclusion

I have attempted to show here how the effects of film music on meaning, memory, and the construction of a reality within a film can be addressed from a cognitive perspective. A Congruence-Associationist framework has been proposed as a basis for new research. In particular, I have hypothesized that conjunctions of musically generated affective associations and film meanings account for a sense of reality or suspension of disbelief created by films. I have also suggested that music draws upon a pool of resources, that, at some level, is independent of resources needed for processing and remembering narrative, speech, and visual aspects of the film.

The proposed framework accommodates formal (structural) and associationist (meaning) aspects. The framework enables systematic exploration of effects of music on visual meaning, memory, and belief. It takes into account bottom-up and top-down processing, typically used in accounts of cognitive processing within a single domain (such as Narmour 1991), here extending that theory to several domains simultaneously. The framework also incorporates constraints of memory. This framework will perhaps facilitate cognitive research on the role of music in cinema and assist in building bridges among cognitive researchers and film-music scholars in the humanities. It may elucidate the awesome cognitive capability of film spectators who make sense of an artful barrage of auditory and visual information—a barrage that in only limited ways matches the patterns of stimulation to which we are accustomed in the real world.

Notes

This work was supported by research grants from the Social Science and Humanities Council of Canada (SSHRC) as well as a SSHRC Canada Research Fellowship. Appreciation is expressed to the editors of this volume and to Charlene Morton for suggestions on an earlier version of this essay.

1. This procedure was invented by Osgood, Suci, and Tannenbaum (1957). It uses many bimodal adjective pairs (like happy-sad, hot-cold, nice-awful) in order to obtain affective meaning on three basic independent dimensions, Evaluation, Potency, and Activity. Osgood's first studies of meaning concerned auditory stimuli, and he was very interested in synaesthetic experience. Tannenbaum (1956) also published a study on the role of background music in live drama and television performance.

2. Using separate groups for each excerpt of interest avoids the problem of bias from prior exposure.

3. Such inattention to music extends to other media involving image-music relationships. In another study (Cohen 1987), students watched a slide presentation about their university campus. The presentation had one of four backgrounds: jazz, appropriate, inappropriate, or none. All subjects were asked to rate the appropriateness of the musical soundtrack. Only one of twenty-four subjects in the condition that had no soundtrack commented on its absence. All others indicated their appreciation of this nonexistent background music; their average fell between that for appropriate and inappropriate music.

4. As David Raskin has written of film music, "its great usefulness is the way in which it performs its role without an intervening conscious act of perception" (in Burt 1994, 5). John Huston credits Alex North with being able to "convey an emotion to the audience with its hardly being aware of the existence of a score" (in Karlin 1994, 204). Inaudibility is of course a major claim of Gorbman's (1987).

5. Gardner's evidence for the intelligences is based on specialized neural centers for these intelligences, a unique developmental course, examples of precociousness as well as isolated presence in savants, cross-cultural universals in the display of these abilities, and distinct notation systems that have emerged for each.

6. The framework extends the Congruence-Associationist framework introduced by Marshall and Cohen (1988) and clarified by Bolivar et al. (1994). A model of attentional focus that considers the interaction of accent structure and association has also been proposed by Lipscomb and Kendall (1994).

References

Benjamin, W. (1969). "The Work of Art in the Age of Mechanical Reproduction." In *Illuminations,* ed. H. Arendt, trans. M. Zohn (219–53). New York: Schocken.

Bolivar, V. J., A. J. Cohen, and J. C. Fentress. (1994). "Semantic and Formal Congruency in Music and Motion Pictures: Effects on the Interpretation of Visual Action." *Psychomusicology* 13:28–59.

Boltz, M., M. Schulkind, and S. Kantra. (1991). "Effects of Background Music on Remembering of Filmed Events." *Memory and Cognition* 19:595–606.

Bregman, A. (1990). *Auditory Scene Analysis.* Cambridge, Mass.: MIT Press.

Bregman, A. S. (1993). "Auditory Scene Analysis: Hearing in Complex Environments." In *Thinking in Sound: The Cognitive Psychology of Human Audition,* ed. S. McAdams and E. Bigand, 10–36. New York: Oxford University Press.

Brown, R. (1994). *Overtones and Undertones: Reading Film Music.* Berkeley and Los Angeles: University of California Press.

Burt, G. (1994). *The Art of Film Music.* Boston: Northeastern University Press.

Chion, M. (1990). *Audio-Vision: Sound on Screen.* Ed. and trans. C. Gorbman. (1994). New York: Columbia University Press.

Cohen, A. J. (1987). "Effects of Musical Soundtracks on Attitudes toward and Memory for Information in a Slide Presentation." Joint meeting of Canadian

Society for Brain, Behavior and Cognitive Science and the British Experimental Psychological Society, Oxford, England.
——. (1990). "Understanding Musical Soundtracks." *Empirical Studies of the Arts* 8:111–24.
——. (1993). "Associationism and Musical Soundtrack Phenomena." *Contemporary Music Review* 9:163–78.
——.(1994). "Introduction to the Special Volume on the Psychology of Film Music." *Psychomusicology* 13:2–8.
——. (1995). "One-Trial Memory Integration of Music and Film: A Direct Test." Paper presented at the annual Meeting of the Canadian Acoustical Association, Quebec.
——(Unpublished). "Processing of Music and Film: Independence and Integration. Manuscript in progress.
Crowder, R. G. (1993). "Auditory Memory." In *Thinking in Sound: The Cognitive Psychology of Human Audition,* ed. S. McAdams and E. Bigand. 113–45. New York: Oxford University Press.
Crowder, R. G., M. L. Serafine, and B. Repp. (1990). "Physical Integration and Association by Contiguity in Memory for Words and Melodies of Songs." *Memory and Cognition* 18:469–76.
Deutsch, D. (1970). "Tones and Numbers: Specificity of Interference in Short-Term Memory." *Science* 168:1604–5.
Dingwall, W. O. (1993). "The Biological Basis of Human Communicative Behavior." In *Psycholinguistics,* ed. J. B. Gleason and N. B. Ratner, 40–48. New York: Harcourt, Brace, Jovanovich.
Fodor, J. (1983). *Modularity of Mind.* Cambridge, Mass.: MIT Press.
Gardner, H. (1983/1993). *Frames of Mind: The Theory of Multiple Intelligences.* New York: Basic Books.
Gorbman, C. (1987). *Unheard Melodies: Narrative Film Music.* Bloomington: Indiana University Press.
Grossberg, S. (1995). "The Attentive Brain." *American Scientist* 83:438–49.
Jourdain, R. (1997). *Music, the Brain and Ecstasy: How Music Captures Our Imagination.* New York: Morrow.
Karlin, F. (1994). *Listening to the Movies.* New York: Schirmer.
Kester, D. B., A. J. Saykin, M. R. Sperling, M. J. O'Connor, L. J. Robinson, and R. C. Gur. (1991). "Acute Effect of Anterior Temporal Lobectomy on Musical Processing." *Neuropsychologia* 29:703–8.
Kimura, D. (1964). "Left-Right Differences in the Perception of Melodies." *Quarterly Journal of Experimental Psychology* 16:355–58.
——. (1967). "Functional Asymmetry of the Brain in Dichotic Listening." *Cortex* 3:163–78.
Lipscomb, S., and R. Kendall. (1994). "Perceptual Judgment of the Relationship Between Musical and Visual Components in Film." *Psychomusicology* 13:60–98.
Marshall, S., and A. J. Cohen. (1988). "Effects of Musical Soundtracks on Attitudes to Geometric Figures." *Music Perception* 6:95–112.
Milner, B. (1962). "Laterality Effects in Audition." In *Interhemispheric Relations and Cerebral Dominance,* ed. V. B. Mountcastle, 177–95. Baltimore: Johns Hopkins University Press.
Mitroff, I., and W. Bennis. (1993). *The Unreality Industry.* New York: Oxford University Press.
Narmour, E. (1991). "The Top-Down and Bottom-Up Systems of Musical Implication: Building on Meyer's Theory of Emotional Syntax." *Music Perception* 9:1–26.

Neisser, U., and R. Becklen. (1975). "Selective Looking: Attending To Visually Significant Events." *Cognitive Psychology* 7:480–94.
Osgood, C. E., G. J. Suci, and P. H. Tannenbaum. (1957). *The Measurement of Meaning.* Urbana: University of Illinois Press.
Patel, A. D., and I. Peretz. (1997). "Is Music Autonomous from Language? A Neuropsychological Appraisal." In *Perception and Cognition of Music,* ed. I. Deliege and J. Sloboda, 191–215. Hove, U.K.: Psychology Press.
Peretz, I. (1993). "Auditory Agnosia: A Functional Analysis." In *Thinking in Sound: The Cognitive Psychology of Human Audition,* ed. S. McAdams, and E. Bigand, 199–230. New York: Oxford University Press.
Samson, S., and R. Zatorre. (1991). "Recognition Memory for Text and Melody of Songs after Unilateral Temporal Lobe Lesions: Evidence for Dual Encoding." *Journal of Experimental Psychology: Learning, Memory and Cognition* 17:793–804.
——. (1994). "Contribution of the Right Temporal Lobe to Musical Timbre Discrimination." *Neuropsychologia* 32:231–40.
Stein, B. S., and J. D. Bransford. (1979). "Constraints on Effective Elaboration: Effects of Precision and Subject Generation." *Journal of Verbal Learning and Verbal Behavior* 18:769–77.
Tannenbaum, P. (1956). "Music Background in the Judgment of Stage and Television Drama." *Audio-visual Communication Review* 4:92–101.
Thompson, W. F., F. A. Russo, and D. Sinclair. (1994). "Effects of Underscoring on the Perception of Closure in Filmed Events." *Psychomusicology* 13:9–27.
Trehub, S. E., A. J. Cohen, and L. Guerriero. (1985). "Development of Emotional Sensitivity To Music." *Proceedings of the 12th International Congress on Acoustics,* Toronto, 3:K5–4.
Treisman, A., and H. Schmidt. (1982). "Illusory Conjunctions in the Perception of Objects." *Cognitive Psychology* 14:107–41.
Zatorre, R. J., A. C. Evans, and E. Meyer (1994). "Neural Mechanisms Underlying Melodic Perception and Memory For Pitch." *Journal of Neuroscience* 14:1908–19.
Zatorre, R. J., and A. R. Halpern (1993). "Effect of Unilateral Temporal-Lobe Excision on Perception and Imagery of Songs." *Neuropsychologia* 31:221–32.
Zatorre, R. J., A. R. Halpern, D. W. Perry, E. Meyer, and A. C. Evans. (1996). "Hearing in the Mind's Ear: A PET Investigation of Musical Imagery and Perception." *Journal of Cognitive Neuroscience* 8:29–46.