ICMPC SEOUL 1998

Proceedings of the Fifth International Conference on MUSIC PERCEPTION AND COGNITION

August 26-30, 1998 Seoul, Korea

The Functions of Music in Multimedia: A Cognitive Approach.

ANNABEL J. COHEN
Department of Psychology, University of Prince Edward Island
Charlottetown, P.E. C1A 4P3
acohen@upei.ca

Music in multimedia abounds with countless examples from television, video, course ware, documentaries, computer and electronic games, professional presentations, advertisements and voice messaging. Technological advances account in large part for this prevalence. Yet, defined as the integration of two or more media types (Roads, 1996, p. 778), or "the simultaneous use of data in different media forms" (Buford, 1994, p. 2), clearly multimedia has had a long history in the form of drama, dance and song, and in these contexts music has always played a part. Music is of course an art form, and its presence might be explained by the aesthetic value alone. But, as will be argued, there are at least seven other functions of music in multimedia contexts.

Whereas multimedia may be informative, didactic, or entertaining, in the first instance these effects take place in the mind. Multimedia "works" when the multisensory information is encoded, interpreted, and stored such that the information can be acted upon. The entire process falls within the scope of cognitive psychology, as does the specific problem at hand of how music works within the multimedia context.

Although phenomenologically, the perception of music occurs immediately and passively, over 20 years of research reveal the complex processes entailed as the brain represents musical sounds, groups them, relates them to other similar patterns, engenders emotional meanings from them etc. With the extent of cognitive processing involved, one might wonder whether music detracts from other aspects of multimedia processing. Recent neurophysiological and psychological research suggests that music activates independent brain functions that are separable from verbal and visual domains. This notion of independence forms a basic assumption of this paper. We now consider the separate functions that the musical mental processes accomplish.

1. Masking. Music was an essential component of the silent film. One of its functions was

regarded as mundane: drowning out the sound of the film projector. Music masked unwanted sound. Masking is defined as "the process by which the threshold of audibility for one sound is raised by the presence of another (masking) sound" (Moore, 1989, p. 84). Similarly, today, music masks distractions produced by the multimedia machinery (hum of disk drive, fan, motor etc.), or sounds made by people, as multimedia often occurs in social or public environments. Music masks unwanted noise in three ways. First, psychoacoustics research reveals that low sounds mask higher sounds. Music is thus an effective mask because musical tones occupy the lower frequencies (fundamentals to 5000 hz but typically below 1000 Hz), as compared to many noise sources. Second, Bregman (1993) and Gaver (1993) point out that listeners aim to hear coherent events in their environment. As music is coherent it will occupy limited auditory attention that would otherwise have been dissipated on extraneous sounds. Finally, sounds embedded in a sequence of sounds are more difficult to detect than sounds in isolation. This is known as informational masking (Kidd & Watson, 1992). In this case, somewhat independent of frequency overlap or range, the presence of music makes other sounds more difficult to hear.

A downside to the masking effect of music is the masking of wanted sounds such as speech due both to overlap in range and the masking of higher speech sounds by lower musical sounds. Thus, filmscore writers must carefully orchestrate background for sections of a film having conversation (Rosza, 1982, p. 26). Sound recording engineers deal with this problem via attenuation of relevant frequency bands.

2. Provision of continuity. Music is sound organized in time and this organization helps to connect disparate events in other domains. Thus, a break in the music can signal a change in the narrative or, conversely, continuous music signals the continuation of the current theme. The music is used as a kind of glue. The capacity-limited brain is forced to "go with the flow". It seems as if the most prominent musical gestalt attracts the concurrent media events from different domains. The role of continuing or breaking the action is a standard technique of film music.

3. Direction of attention. My collaborators and I have previously argued that structural aspects of music may direct attention to specific aspects of a visual scene (Marshall & Cohen, 1988; Bolivar, et al., 1994). Visual and auditory stimuli can be described on two dimensions: meaning and structural features. Structural features refer to such aspects as rhythm, contour, or patterning in time. For example, the temporal structure of an object expanding or contracting three times a second can match a sound that increases and decreases in intensity three times a second. In this example, the formal properties of visual objects are congruent with those of the music. The relation is not necessary, but, it is proposed that when this structural congruence occurs it has special significance to the viewer-listener. This reasoning follows from theories about the significance of formal congruence *within* a modality. For example, Bregman (1993) discusses auditory scene analysis as dependent on properties described by the Gestalt psychologists as similarity, proximity, and good continuation. In trying to make sense of the world, the brain defines objects, and objects are those things that produce correlated sensory patterns. For example, speech formants having the same amplitude modulation pattern likely come from the

same speaker, and therefore are grouped together by the listener as the speech from an individual speaker.

Following this argument, patterns from different modalities that have the same pattern of change across time are likely to reflect a single object. Thus, when auditory patterns mimic visual patterns, attention would be directed to that visual pattern rather than one which is not congruent with the auditory background. The detection of auditory visual congruence finds support in the work of Lipscomb (1996) who shifted the auditory accompaniment of three kinds of video material differing in complexity. The shift was detectable and affected the influence of the background music on the two simpler patterns, though not the most complex. Bolivar et al., (1994) also showed that viewer/listeners agree on what visual and musical patterns are temporally congruent.

4. Mood induction. Music can alter the mood or feeling of the listener-viewer. Evidence for the mood-altering ability of music comes from a variety of sources. Pignatiello, Camp and Rasar (1986) selected excerpts of music that they felt were specifically sorrowful, happy, or neutral. Compilations of the three types of music were produced. Following the presentation of a tape, the subjects subsequently completed cards about how they were feeling. It was clear from these judgments that the music had altered their moods in the expected direction. This technique has also been used by Albersnagel (1988) to stimulate depression and elation. Therefore, music can be used to encourage positive multimedia experiences.

It is important to distinguish between mood induction and communication of meaning by music. Mood induction changes how one is feeling, while communication of meaning simply conveys information. One may receive information depicting sadness without him or herself feeling sad (Rosar, 1994). It is possible that such musical meaning may elicit associated feelings, but it is not necessarily the case.

5. Communication of meaning. The visual screen is often referred to as two-dimensional (see Palmer, 1990; Rosar, 1994) with music adding emotion as a third dimension. In the days of the silent films, an entire industry surrounding background musical performance emerged (Limbaucher, 1974). Books provided lists of songs capable of conveying certain emotions, and songs were categorized with respect to the kind of emotion each communicated.

Different features of music convey emotional meaning. For example, sadness is conveyed by slow pace, falling contour, low pitch, and the minor mode and happiness is conveyed by fast tempo, rising tempo, high pitch and the major mode (Levi, 1982; Trehub, Cohen & Guerriero, 1985). Music can take on meanings through association with events (Cohen, 1993); e.g., music at funerals becomes associated with sadness. Communication of emotional meaning or meaning through association (e.g., of epochs, cultures, and events) is particularly effective in situations that are ambiguous. For example, as described by Cohen (1993) two contrasting excerpts of background music systematically altered the interpretation of two people either fighting or playing Yet for an unambiguous fight scene, these same backgrounds pieces of music had little effect. Sirius and Clarke (1994) showed that background music changed the meaning of a simple animation of a moving geometric figure,

and Bolivar et al., (1994) showed that aggressive and friendly music could systematically alter the friendliness or aggressiveness of wolves in social interaction.

Composer Marvin Hamlich, interviewed on television, has said that for a composer, underscoring is like simultaneous translation, except rather than translating into French or Swahili, the second language is that of music. The film-score composer has a practical knowledge of the semantics of music tacitly known to the majority of viewer-listeners. Iwamiya (1994) also suggests that producers of music backgrounds for multimedia have implicit knowledge about music and video meanings. He compared two types of video and music background combinations. The music and video were either presented together as "packaged" or modified by audio-video switching or audio-video desynchronizing. Viewer-listeners gave lower ratings of aesthetic judgments to the altered versions. This suggests that producers of the multimedia productions and the audience are sensitive to the same variables of meaning and congruence.

6. Music as a cue for memory. Music can take on the meaning of that which it accompanies. This is the rationale of the technique in opera known as leitmotiv (Gorbman, 1987; Gregory, 1995). Here it is assumed that a musical motiv can come to mean a particular individual, or theme, such that in the absence of the individual or other visual cues, the music alone will cause people to imagine the individual or theme. There is a psychological basis for such a mechanism in the classically (Pavlovian) conditioned response (cf., Cohen, 1993). Theorists of film music sometimes suggest that these music-meaning associations arise in just one presentation and this has received some empirical support. For example, Boltz, Schulkind and Kantra (1991) showed that recognition of a single visual frame from a previously viewed television program was enhanced if the music that had accompanied that portion of the program was played during the presentation of the to-be-recognized material. Cohen (1995) also showed that such connections can form after one presentation. In this study, 20 short video and music excerpts were presented simultaneously. Subsequently, a series of music-video excerpts were presented. Five of the excerpts were paired as in the original, while for the others, groups of five had either new audio, new video, were mismatched, or were completely new. The subjects rated how sure the music-video test pair had been presented together. The ratings were higher for the matched than the mismatched pairs, showing that within just one exposure, links between the music and visual modalities can form in short-term memory (STM).

The use of music as a memorial cue is a function capitalized on by advertisers. The advertiser hopes that a musical "hook" will be associated with the product. The music may also signal a particular television program, such as the evening news, or the start-up of a computer operating system such as Windows-95. The music can thus be prepare the mind for a type of cognitive activity.

7. Arousal and focal attention. Given the model of a mind that has separate places for components of music, it is a simple fact that when there is music, more of the brain is active. Increased activity of the brain may increase concentration on the primary attentional focus and filter out distractions such as the frame around a monitor or screen, people in the vicinity, or the

pressure from one's chair. Heightening arousal level may thus increase the intended impact of the multimedia experience, making the experience seem real. It can be suggested that heightened arousal temporarily suspends contact with "real" reality and facilitates belief in the virtual reality. However, there is evidence that excess stimulation alters attention towards audition and peripheral vision and away from central vision (Shapiro & Lim, 1989).

8. Musical aesthetics. As mentioned at the outset, music is an art form and its presence enhances every situation in much the same way that a beautiful environment enhances the experience of activities within in. Little space will be devoted here to this aspect of music's role in multimedia, but the aesthetic function of music should never be forgotten. Of course, music that is not appealing can disturb the user, and certain personality characteristics may determine whether an individual can or cannot cope with background music (e.g., introverts may need and want less stimulation, so that musical background would not enhance the multimedia experience at this level).

A congruence-associationist flow diagram as a framework for understanding the

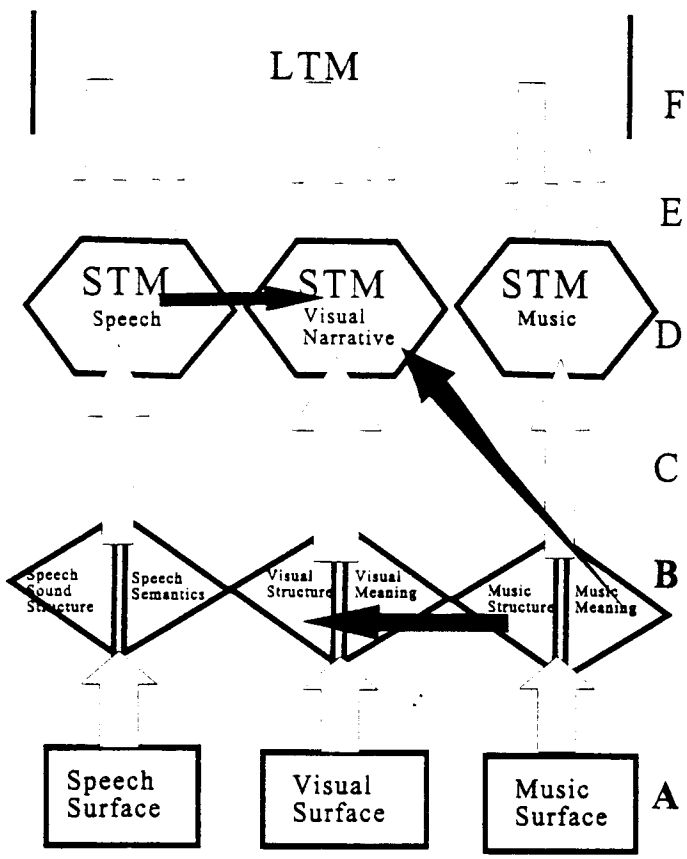

Fig. 1. **Congruence-Associationist Flow Diagram**

function of music in multimedia contexts. As reviewed above, music serves at least eight different functions. It masks distraction, provides continuity, directs attention, induces mood,

communicates meaning, cues memory, heightens arousal and suspends disbelief, and finally adds an aesthetic dimension. Figure 1 provides a flow diagram that can be useful in examining how these various functions co-exist in a multimedia context. The framework is referred to in terms of congruence and association in order to capture the structural and meaning components of the three continuous media involved: music, video and speech (see also Cohen, in press). One example of the use of the diagram is developed below.

Level A represents the physical surface of speech, video and music. At level B, each of these is decomposed into structural and meaning components. For music, this means decomposition into temporal structures and emotional and associative meanings. At level C, attention is especially paid to that part of the visual scene that shares temporal structural features with the music, from the grey oval represented at B. This part of the visual information is depicted at D where the emotional information from B is also sent. Thus, in STM, the visual scene is recombined with the emotional meaning, but leaves the acoustical aspects of the music as neglected residue. In the context of his Adaptive Resonance Theory of conscious attention (ART), Grossberg (1995) has argued that material in STM comes into attentional consciousness if it is matched by material from LTM. This explains why the acoustical aspects of the music would not be attended. The music does not make sense to LTM (level F) (where is that background music coming from?) and no hypotheses can be generated to include it. So the phenomenal experience (level E) is one of a narrative with visual and emotional components. At some other part of consciousness, the acoustical aspects of music can be processed, as it is known that simultaneous tasks can co-occur (Neisser & Becklin, 1975) and there is evidence that this music is encoded in memory (Boltz et al., 1991; Cohen & Dunphy, 1990). generated to include it. et al., 1991; Cohen & Dunphy, 1990). generated to include it.

This flow diagram emphasizes that music is a vehicle transporting a variety of information that serve various multimedia goals. The brain selects what is useful. A prime example of this is the role that emotional meaning from music provides to a visual narrative while the sounds of the music are of secondary concern. Yet, as a signal that it is time to watch the news, it is the sound pattern itself that is important.

Thus, the present paper has drawn attention to the multidimensionality of music, and the modular and interactive aspects of music processes, as both an explanation and highlighting of the consequent multiple functions of music in multimedia contexts.

References

Albersnagel Velten and musical mood induction procedures: A comparison with accessibility of thought associations. Behavior Research Theory, 26, 79-96.

Bolivar, V. J., Cohen, A. J., & Fentress, J. C. (1994). Semantic and formal congruency in music and motion pictures: Effects on the interpretation of visual action. Psychomusicology, 13 (1), 28-59.

Boltz, M., Schulkind, M., & Kantra, S. (1991). Effects of background music on remembering of filmed events. Memory and Cognition, 19, 595-606.

Bregman, A. S. (1993). Auditory scene analysis: Hearing in complex environments. In S. McAdams, & E. Bigand (Eds.), Thinking in sound (pp.10-36). New York: Oxford

Buford, J. F. (1994) Uses of multimedia information. In J. F. Buford (Ed). Multimedia systems. (pp. 1-25). Reading MA. ACM Press. Addison Wesley.

Cohen, A. J. (1990). Understanding musical soundtracks. Empirical Studies of the Arts, 8, 111-124.

Cohen, A. J. (1993). Associationism and musical soundtrack phenomena. Contemporary Music Review, 9, 163-178.

Cohen, A. J. (1994). Introduction to the special volume on the psychology of film music. Psychomusicology, 13, 2-8.

Cohen, A. J. (In press). Film music: A cognitive perspective. In D. Neumeyer, J. Buhler, & C. Flinn (Eds.), Music and Cinema. Wesleyan Univ. Press.

Cohen, A.J and Dunphy, D. (1990). Musical and visual processing in film. Canadian Psychologist, 31, 220 (Abstract). Poster Annual meeting Canadian Psych. Assoc. Ottawa.

Gaver, W. W. (1993). What in the world do we hear? An ecological approach to auditory event perception. Ecological Psychology, 5, 1-29.

Gorbman, C. (1987). Unheard melodies. Bloomington: Indiana University Press.

Gregory, A. (1995). Recognition of Wagner's Leitmotives. Society of Music Perception and Cognition Conference. Berkeley. p. 73. abst.

Grossberg, S. (1995). The attentive brain. American Scientist, 83, 438-449.

Iwamiya, S. (1994). Interaction between auditory and visual processing when listening to music in an audio visual context. Psychomusicology, 13, 133-153.

Kidd, G. R. & Watson, C. S. (1992). The "proportion of the total duration rule" for the discrimination of auditory patterns. Journal Acoustical Society of America, 92, 3109-3318.

Levi, D.S. (1982). The structural determinants of melodic expressive properties. Journal of Phenomenological Psychology, 13, 19-44.

Limbacher, J. L. (1974). Film music. Metuchen: Scarecrow.

Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually significant events. Cognitive Psychology, 7, 480-494.

Lipscomb, S. (1996). Cognition of musical and visual accent structure alignment in film and animation. Proceedings of the 4th ICMPC Montreal. pp 309-313.

Marshall, S. & Cohen, A. J. (1988). Effects of musical soundtracks on attitudes toward animated geometric figures. Music Perception, 6, 95-112.

Moore, B. (1989). Introduction to the psychology of hearing. 3rd Ed. N.Y.: Academic.

Neisser, U. & Becklen, r. (1975). Selective looking: Attending to visually significant events. Cognitive Psychology, 7, 480-494.

Palmer, C. (1990). The composer in Hollywood. N.Y.: Marion Boyars.

Pignatiello, M. F., Camp, C. J., & Rasar, L. (1986). Musical mood induction: An alternative to the Velten technique. Journal of Abnormal Psychology, 95, 295-297.

Roads, C. (1996). A tutorial in computer music. Cambridge: MIT Press.

Rosar, W. (1994). Film music and Heinz Werner's theory of physionomic perception. Psychomusicology, 13, 154-165.

Rozsa, M. (1982). Double life. New York: Hippocrene Books, Inc.

Shapiro, K. L., & Lim, A. (1989). The impact of anxiety on visual attention to central and peripheral events. Behavior Research Therapy, 27 (4), 345-351.

Sirius, G. & Clarke, E. (1994). The perception of audiovisual relationships: A preliminary study. Psychomusicology, 13, 133-153.

Thompson, W. F., Russo, F. A., & Sinclair, D. (1994). Effects of underscoring on the perception of closure in filmed events. Psychomusicology, 13, 9-27.

Trehub, S. E., Cohen, A. J. & Guerriero, L. (1985). Development of emotional sensitivity to music. Proc. 12th International Congress on Acoustics, Toronto, 3, K5-4.

Author Note

The work was supported by grants from the Social Sciences and Humanities Research Council (SSHRC).