Musikpsychologie Bd. 18, 11–21, © Hogrefe Verlag, Göttingen 2006

Themenschwerpunkt
Inter- und Multimodale Wahrnehmung

Audio-visual perception and its relevance in science and art

Helga de la Motte-Haber

Zusammenfassung

Die alltägliche Wahrnehmung basiert immer auf multisensorischer Information. Dennoch hat sich die Wissenschaft auf die Verarbeitung der Information aus einem Sinnesorgan konzentriert. Auch in der Kunst fand eine solche Trennung der Sinnesorgane statt. In jüngerer Zeit jedoch konzentrieren sich Künstler sehr stark auf multimediale Formen. Auch die wissenschaftlichen Untersuchungen beschäftigen sich immer mehr mit der Zusammenarbeit der Sinne. Der Aufsatz will grundlegende Untersuchungen zusammenstellen und Parallelen zu neuen Kunstformen ziehen. Der weit verbreiteten Auffassung der Dominanz des Auges wird auf Grund von experimentellen Befunden widersprochen.

Abstract

Normally, daily life-perception is based on multi-sensory information. Nevertheless information processing is only studied for single modalities. There is a comparable separation in traditional art which normally refers only to one sensory modality. Recently as well in science as in art production the situation has changed. The paper intends a summary on audiovisual perception and its relevance for art. Results on audiovisual perception are presented which contradict the widespread conviction that the eye is the dominant sensory modality.

1. Introduction

In the following paper I will outline some basic information on audio-visual perception. It will be shown that different sensory modalities are related to different information which mutually completes in a holistic way our daily

life perception. Normally, however, a paradigm with conflicting visual and auditory stimuli has been used for psychological research. The main idea is not to show this holistic (global) functioning of the information processing but to determine the degree of contribution of the different sensory modalities. The widespread opinion of psychologists that the eye is the dominant modality will be proven critically.

In this paper the phenomenon of cross modality-equivalence will be discussed. Synaesthesia is only one form of cross modality equivalence. Therefore, cross modality-equivalence and synaesthesia are treated separately. Only a few references will be given concerning the development of audio-visual perception during infancy.

My involvement in audio-visual perception results not only from pure psychological interest. It results mainly from the importance of aesthetic in general but also from the particular significance on contemporary art. I will therefore start with a preface pointing out the aesthetic premises.

2. Preface

(1) Traditionally, the arts are separated into disciplines delimited by sensory modalities. We are seldom aware that this separation did not occur prior to the 18th century. In the olden times, music, painting, sculpture, and poetry belong together. They were unified by a common religious or secular function and belong to specific but common places like a church, a castle, or a market. As the idea of independent work increased, the arts lost their functionality which led them to combine themself into special occasions. Special institutions then became necessary. The new aesthetic has also been accompanied by the creation of separate specific venues: the concert hall or the museum.

(2) The idea of an anti-functional autonomous artwork has been rooted in the strong desire of the artists to be free from any obligation, particularly from any aristocratic directives. Therefore, it is possible to say that the revolution of the middle-class and their claim to freedom is based on the presupposition of the present aesthetic conceptions. Sotto voce: Background music is excluded from this statement.

(3) But how does one substantiate the liberation of an artist? How can it be possible to distinguish his creation from a purely subjective arbitrary act? It was not sufficient to claim freedom from functions or commissions which only brought about a lack of money. Therefore, further arguments were necessary which would show the artist's obligation based on facts. The idea that he is bound to a specific material has been helpful to justify the independence of both the artist and his work. This means: the painter is only bound by pencil and colour; as his duty is to be a painter. For the musician, similar binding circumstances result from sound. This commitment has been the basic argument for the separation of the arts

requiring different forms of perception through the eyes or the ears. My short preface shows an abbreviate summary of the famous essay *Laokoon* by Gotthold Ephraim Lessing, published in 1766. For the artists, it has the "impact of a drug" as is shown in Johann Wolfgang von Goethe's *Dichtung und Wahrheit.*

(4) The doctrine of the independent artwork lasts until today. But soon it didn't need an explanation. It has become self-evident. No further proof was necessary. In addition to this, the demand on delimited access by different senses seemed to be superfluous. One of the dreams of the romantic tradition has been that works particular to each artistic discipline might be meaningfully represented in another artistic discipline. The challenge on the inter-disciplinary translation of artworks results in two forms of solution: Artists were seeking structural analogies. At the beginning of the 20th century, abstract rhythmic analogies in particular offered a comfortable possibility to compare music with painting in anticipation of the latter's abstract forms. For a quick illustration I will mention the *Fuge in Rot* by Paul Klee. Beyond the 19th century artists already dared the combination of music and painting.

For example, Caspar David Friedrich advised the buyer of his paintings to view them during the dusk and with accompanying music. During the last 100 years, many artists – Wassily Kandinsky, Kurt Schwitters, the young Arnold Schönberg, John Cage, Morton Feldman, or Karlheinz Stockhausen – devoted their works to the challenge of overcoming the separation of the arts. Very often the idea of the "Gesamtkunstwerk" by Richard Wagner still has had a stimulating power, though in the 20th century artists have created new forms. The belief in artworks separated into disciplines delimited by different sensory senses no longer exists. More and more we become aware that works have never been realized in its strictest sense. Normally, listening to music in a concert is based on an audio-visual information processing. It is transmitted by visual stimuli, which enhance the listening process. Therefore, recorded music without visual impact creates problems which are corrected by technical compensation. In the 20th century the visual appearance of music became a fascinating challenge to some composers as demonstrated by the "Instrumental theatre" of Mauricio Kagel. Later, new forms of entertainment (i. e. multimedia shows) have also aroused.

Normally, books on psychology of music lack chapters on audio-visual information processing. If there is an interest by a psychologist in audio-visual perception, it only concerns the emotional impact of music on a film. Some researches are conducted based on sight reading and in rare cases, on synaesthesia. With regard to general aesthetic development, a broader knowledge seems to be necessary. Today we have new multimedia forms. The latest genres are the internet art and the sound art, which both require additional careful studies on audio-visual perception.

3. Basic researches on the interaction between eye and ear

3.1 Dominance of the eye for spatial qualities

General studies on information processing simultaneously produced by eyes and ears are often based on the general idea that the eye is the dominant sensory modality. Many studies on the effect of ventriloquism show the shifting of a sounding source in the direction of a supposed visual source. This interesting circus effect can also be experienced in cinema. It seems that the actors speak directly from the screen though the loudspeakers are on the left and right sides. The shifting effect works well in the horizontal plan if the sound source is hidden. Several authors found an angular limit in the horizontal median plan to be around 40 degrees, whereas much wider angles have been shown in the vertical plan (Thurlow & Jack 1973, Mateeff, Hohnsbein & Noack 1985). Some doubts arise about the idea of a general dominance of the eye by basic experiments which use only abstract lights and sounds. Similar to the effect of ventriloquism, a shift in the sound source has been noted if a light is moving in a space although the shifting is delimited by a smaller angular of 15 degrees (Radeau 1994). Contrary to the circus effect, Radeau shows also a shift of the light into the direction of the sound source. However, this shift of the light is minimal compared to the shift of the sound.

Another experiment which demonstrates the dominance of the eye concerns the estimation of distances. I will mention an experiment undertaken during a conference of the *Acoustical Society of America* (Gardner 1968). Several loudspeakers were lined up in a row before a person. Only the first loudspeaker is visible while the rest behind it are indicated by numbers above them. Even if the sound came from a loudspeaker nine meters away, it was actually detected from the visible loudspeaker which is directly in front of the person. Other participants of this conference viewing from the side should have laughed. This experiment seems to be very abstract but it has relevance on the position of the loudspeakers during a performance of electro-acoustic music. Some of you may have already seen the loudspeaker orchestra of the *GRM* in Paris with which composers like François Bayle create moving spaces. For them, a careful spatial layout of the loudspeakers is necessary. Instead of ventriloquism, it is better to speak of a visual capture effect. This more general term is also valid for a temporal shift of sound which is of special interest to the multimedia world. Very seldom during a television-music-transmission, sound and pictures are totally synchronous. Today, by technical reasons it is normal that music video clips have a delayed sound.

There are several experimental studies which demonstrate this effect of temporal synchronism even if the sound is delayed for about 40–100 milliseconds. I will mention one elegant design. Cavé, Ragot & Fano (1992) showed a film with one visual event: A film flap folded up. During several presentations the sound has been successively delayed. Guessing how good the synchronization of the visual and the auditory dubbing was, the subjects have been tolerant until a delay difference of 100 milliseconds. But the ex-

periment also shows a total intolerance of the subjects if the sound came too early. As a musicologist more interested in the ear than the eye, I know that the ear needs only two milliseconds for the registration of a temporal difference contrary to the eye with a lower ability of 17 milliseconds. Therefore, I did not completely believe in the interpretations of the results with regard to the general dominance of the eye.

We made several experiments in the university on time delay of sound. For example, I showed a film for two minutes about a walking and running woman whose feet are visible. Another experiment concerns the film *Pas de Cinq* by Maurico Kagel, where the actors have to walk rhythmically. It has been confirmed by all these experiments that a strict synchronization of sound and pictures is not important. However the question arises as to why experimental filmmakers endeavour so much to synchronize the different layers. Even Oskar Fischinger strained every nerve in the 1930s for that synchronization as well as his followers the famous artists from the American West Coast, like John & James Withney, Jordan Belson and Norman McLaren.

Recently we were surprised by the results of an experimental design using the simplest material: computer-generated and coordinated beeps and flashes. The subjects were undergraduate students in musicology. They looked at a flash on a computer screen and listened to a beep by the use of earphones. The beep sounded in advance, synchronous and delay modes. The subjects had to choose only one mode. The results show that nearly 100 % of the subjects discovered correctly the synchronous stimulus. The findings of the above-mentioned experiment of Cavé, Ragot & Fano (about the different impressions of an auditory stimulus in advance and delayed modes) could not be replicated. Also the hypothesis of a general visual capture is not confirmed. Though further research is necessary, it can be presumed that the mutual attachment of asynchronous stimuli is less aroused by modality-specific features and more by a cognitive set, e. g. the semantic meaning of a situation (running women and noise of feet). Semantic factors can influence the multi-sensory pairing (Pourtois & Gelder 2002).

The findings of correct identification of synchronization corresponds to the findings of Scott D. Lipscomb (1995) who investigated accent structure alignment in three abstract animations by Norman McLaren and in excerpts from three normal narrative motion pictures. He created several conditions for the sound: coordination, dissonant and consonant. The range of the shifts were in the size of –890 and +532 milliseconds. Synchronization attracted attention only to the abstract films of Norman McLaren. The different conditions failed to have an influence on narrative film excerpts. Lipscomb concluded that in the case of narrative film excerpts a lack of salient moments in the visual image is responsible for the bad results. I prefer another interpretation stressing the difference between local and global cues. Maybe in the case of the narrative film excerpts neither the direct auditory input nor the visual stimuli were the bases for people's judgment on the story. This means that the miscalculations are not evoked by modality-specific features but by an emotional semantic scheme as a result of a cognitive hypothesis.

Normally, such a scheme or set produce a form of "inattentional blindness" (Simons & Chabris, 1999) in favour of global cues instead of local cues. The data of Lipscomb fit well onto our own findings that a cognitive set can be responsible for a visual capture, – nevertheless, correct judgments of synchronization are possible.

Other investigations show similar effects. For example, the ventriloquism effect doesn't work without an appropriate association between eye and ear. Above all, plausibility is necessary for a quick top-down process. In 1953, an investigation by Jackson revealed that looking at a lamp had little influence on the position of a ringing bell. However, a steam boiler attracts a sniffle within a range of 30 degrees. Let us summarize the psychological research to see if there is any relevance to art.

Daily life perception benefits from the information delivered by multiple modalities. Isolated visual or auditory stimuli very seldom occur. Normally, psychological studies use conflicts to study the interaction between eye and ear. Often they are inspired by the hypothesis of the dominance of the eye. Doubts in a modality-specific dominance are produced by carefully examining the results. These doubts are underlined by investigations showing the balance of the different modalities during childhood (which I have not discussed). Nevertheless, no doubts exist with regard to the findings that in adults perceived location is determined predominantly by visual information. The eye is that sensory modality by which the best judgment is possible when it comes to location and distance. In order to explain a temporal visual capture effect, it is necessary to take learned cognitive schemes into account. Such schemes can produce illusions by simplifying the percept (for example, that a footstep has to be synchronized with the feet). If there is no reason for such a wrong but plausible version, then correct information processing can be possible.

I outlined roughly in my preface the new development of art towards the direction of audiovisual works. Some artists are working with the temporal synchronization effect but psychologists on the contrary do not intend to explain the wrong daily-life-perception by stating the dominance of visual information. Artists aim to stimulate a more general reflection on information processing. Therefore, artists create very often an artificial and strange situation which cannot be surveyed easily. It should be a challenge for the perception because it is impossible to call up the plausibility hypothesis in our mind. Without any help from cognitive automatism we can become aware of the difficulties of information processing. It can be an aesthetic intention to promote the knowledge of our perception as a basic process of our intellectual power. Tilman Küntzel, a sound artist, often uses a complete synchronization of light and sound. Although knowing this, you can have the strange impression of an "auditory capture" effect. Imagine an installation of neon tubes shining rhythmically and accompanied by synchronous sound. After a while the visitor questions why the light seems to be produced by a sound ahead of time. Maybe this example gives rise to the question whether the ear is working faster or why it delivers this kind of experience.

3.2 Dominance of the ear for temporal qualities

Research was largely concerned with identifying separate working processes of the senses. The situation changed in time. Since the 1990s, understanding the interplay of the sensory modalities has become more important. More and more it becomes of interest to show how different modalities provide information about the same external event or object. While earlier researches were largely concerned in showing the dominance of the eye for the orientation, there is an increasing tendency to take into account the differences between the sensory modalities. This shows where, when and which sensory input delivers the best information and how the others contribute. This type of research has been preferred by the pioneering studies of Meredith & Stein (1986) who uncovered special brain areas which were only stimulated by a multi-sensory input. However, a sufficient model for the audio-visual perception could not yet be developed up to the present. There is separate information of various kinds which are suitably weighted as they are combined. Multimodal levels of representation may probably influence unimodal levels through feedback. Driver & Spence (2000) proposed the application of a fuzzy logic. But this is only a proposal as long as studies are no longer conducted. More so, the lack of studies with priority on the influence of auditory stimuli can be ascertained. Why should the ear not be the driver for audio-visual information processing in some situations?

The best information about distances and about the form of objects is read out from visual stimulation giving rise to the wrong conclusion of the dominance of the eye. But with visual information alone we cannot see an empty space. The so called Ganzfeld – studied by Wolfgang Metzger (1930) – can be easily demonstrated by containing the head inside a large evenly illuminated hollow hemisphere. The uniform stimulation creates the impression of a light mist filling an endless space. The well-known artist James Turrell utilizes these findings for his Ganzfeld-Works. (You can find some of them on the internet.)

Though the ear has a tendency to hear sounds at allowable distances, and in accordance with our interpretation of the sounding sources, it is capable to analyze very well the volume of a room. Transformation of spaces by sound is the most frequent form of sound art. Today, it plays also an important role for the soundtrack of films; newly-developed technologies are helpful, for example the Dolby surround, and artificial resonances. Fascinating are the investigations of Shams, Kamitani & Shimojo (2002). They presented that auditory stimuli altered the percept of visual stimuli. If they presented one flash but several beeps, they evoked the impression that there were more flashes. Subjects reported seeing two or more flashes, when a single flash is accompanied by two or more beeps. The beeps follow each other very quickly (with a temporal distance of 57 milliseconds). Therefore, the results reveal that visual illusion induced by sound is evoked only by the frequency of the beeps or by the frequency in time, that is, by the tempo. Musicians know for a long time that audition can dominate vision for temporal properties, because of its

higher acuity. In a movie, it is a common practice to intensify the tempo of pictures by music. Erik Satie already used this method for the surrealistic film *Entr'acte*. He accelerated the pictures at the moment when such have reached its peak.

4. Synaesthesia

Hearing and seeing provides information about the same external object. But sometimes the stimulation of one sensory modality reliably causes a perception in one or more different senses. Thus, listening to a sound can evoke a visual impression, mostly a coloured impression. Usually this phenomenon is described as a form of colour hearing, or as a form of photism. These synaesthetic experiences are involuntary. The photisms appear outside of the body as a projection close to the body. They can be stable through a lifetime where no intersubject constancies exist. I will not enumerate all these different combinations. Only some general characteristics should be mentioned. Colour hearing is predominantly found in women and it occurs in less than one percent of people. An exact matching between pitch and sound presupposes absolute hearing. Olivier Messiaen can be quoted as an outstanding example. The prefaces to his pieces which he calls "musique de couleur" anticipate the description of "coloured-hearing synaesthesia" given by Lawrence E. Marks (1978). For Olivier Messiaen the entire visual field is filled with colour, mostly with several colours simultaneously, that change over time with the music, with each colour reflecting a particular harmonic aspect of the music. Since 1960 these descriptions disappeared slowly in the prefaces of his works because too often he has been confronted with sceptical critics. Therefore we wouldn't completely know about his harmonic thinking as well as his colour system.

Synaesthets never see complex dreams. If figurative photisms are perceived, they will have a geometric shape: impressions like lines, spirals, circles, or lattices have been reported. These percepts bear a considerable likeness to Kluver's "form constants". Heinrich Klüver (later Kluver 1966) showed in his experiments (starting in the1920s) that the influence of mescal produces visual hallucinations with typical geometric shapes. Today it is known that they are similar to the visual impressions of synaesthets. But the process of colour hearing is different with regard to the external auditory stimuli. Tempo for instance affects the shape of the photism: the faster the music, the sharper and more angular the forms. The size of the photism depends on pitch. High-pitched sounds produce small photisms and low-pitched sounds produce visual percepts that are large in size. Loudness has also a direct effect: the louder the bigger.

Synaesthesia has been an important artistic phenomenon in the beginning of the 20th century; it has been used as a veto against the separation of the arts outlined in my introduction. I will mention two outstanding cases of synaesthets: one a painter, the other one a composer. In his famous essay

Wassily Kandinsky had analogized in 1912 the impression of colours with the sonority of instruments (Klangfarbe). He took to his own system of colours the system of Johann Wolfgang von Goethe. He worked out his theoretical thinking twice: first by a plot for a theatre piece *Der gelbe Klang* (1912) and then by an illustration of Modest Mussorgsky's *Pictures of an Exhibition* (1928). The latter is a stage set of moving coloured abstract figures and forms. Kandinsky was convinced that he is a synaesthet. In contrast to Olivier Messiaen, he did not use a stable colour-sound system. For him more or less sound seems to be a metaphor indicating a general symbolic value for all other arts. But Kandinsky foresaw, avant la lettre, a form of music animation.

In the beginning of the 20th century there are many works created to show the relation between different arts by matching sound and colour, for example Arnold Schönberg's *Glückliche Hand*. But no one is as famous as Alexander Skrjabin's *Poème du feu. Promethée*. Skrjabin prescribed in the score a colour-piano written in the usual piano notes. The relation between vision and audition is based on an analogy of the rainbow in circles of fifths. There were several reasons to doubt that Skrjabin, against his own conviction, was a real synaesthet. In particular, the case study of the English psychologist Charles S. Myers published in 1915 called this ability in question. On the other hand, the two-folded moving coloured patterns produced by the two parts of the colour-piano has such a big similarity to Kluver's "form constants". Maybe Skrjabin was in an ecstatic state while composing.

5. Cross-Modality Equivalence

Synaesthesia is no longer a fascinating theme for artists. They are more interested in other forms of cross-modality matching. Instead of colour hearing, musicians are working beyond the boundaries of the conventional genres by concentrating on a-modal features which can be specified in more than one modality. Intensity, duration, rhythm, spatial qualities, shape, and so on are not peculiar to a single modality. They refer to several senses, or to that one which Aristotle called "sensus communis". Lawrence Marks (1978) asked subjects to match pure tones to brightness of grey surfaces. His results indicated that most subjects matched increasing auditory pitch to increasing visual brightness. Marks conducted a series of further experiments in which subjects used scales of loudness, pitch and brightness to evaluate the meanings of various auditory-visual synaesthetic metaphors such as: sound of sunset, murmur of dawn, bright whisper to name a few one. He found that loudness and pitch expressed themselves metaphorically as greater brightness and likewise that brightness expressed itself metaphorically as greater loudness and as higher pitch. In terms of cross-modality matching one might conclude from Marks findings that our senses are integrated somehow. However Stein & Meredith (1993) offer a different point of view based on neurological perspective with the aim to distinguish multi-sensory and inter-sensory phenom-

enons. While cross-modal matching is clearly an inter-sensory phenomenon, and may involve multi-sensory neurons, one could make the case that it has little to do with the integration of inputs from different modalities per se, and that multi-sensory areas of the brain need not to play any special role in this process. The judgments of equivalence across modalities could depend on the individual input being held in the central nervous system in modality-specific form, so that they are independent of one another, but still may be accessed neural pool.

I don't have to go into details to show the artistic relevance of cross-modal equivalence because every musicologist knows that since 1950 composers like Karlheinz Stockhausen, Pierre Boulez, Luigi Nono, Bernd Alois Zimmermann and their followers spread out musicians in the concert halls in order to use spatial stimuli as new material for their compositions. This means they are working with a-modal features valid for several senses. The ability for the cross-modality matching is not yet well-investigated. But it seems to be a basic ability. There are findings which point out that during early infancy, cross-modality equivalence is more important than identifying modality-specific features (Lewkowicz & Turkewitz 1980). Contemporary musicians are working on in new fields. Therefore, for a long time colour hearing has been regarded as old-fashioned.

However today, academic exploration of synaesthesia is again stimulated (Baron-Cohen & Harrison 1996). The modern multi-media-machines exposed a huge field for the application of researches. Presentations of music in clubs or discos (discotheques) are often accompanied by coloured light. They stimulated scientific research in the last years. It is assumed that higher cortical functioning is responsible for synaesthesia. Though this phenomenon occurs in very rare cases, it is interesting to discover the mode of operation of the brain. Neurophysiologists are proud to attest to a synaesthet the stimulation of the visual cortical area while listening to auditory stimuli. The old hypothesis of Bleuler & Lehmann (dating back in 1881) looks through these new investigations.

Ear and Eye work together on several levels of perception. It is necessary to differentiate information processing re inter- and multi-sensory perception and re mutual influence of temporal and spatial qualities. I hope that in the future we will have a better understanding of contemporary art and learn more about the interactions between the different sensory modalities and cross-modal equivalence.

References

Baron-Cohen, Simon & Harrison, John (Eds.) (1996). *Synesthesia: Classic and Contemporary Readings*. Oxford: Blackwell.

Bleuler, Eugen & Lehmann, Karl Bernhard (1881). *Zwangsmäßige Lichtempfindung durch Schall und verwandte Erscheinungen*. Leipzig: Fues's Verlag.

Cavé, C.; Ragot, R. & Fano, M. (1992). Perception of Sound-Image Synchrony in Cinematographic Conditions. In: C. Auxiette, C. Drake & C. Gérard (Eds.), *4. Workshop on Rhythm, Perception & Production* (pp. 25–30). F-Bourges.

Driver, Jon & Spence, Charles (2000). Multisensory Perception: Beyond Modularity and Convergence. *Current Biology, 10 (20)*, 731–735.

Gardner, Mark B. (1968). Proximity Image Effect in Sound Localization. *Journal of the Acoustical Society of America, 43*, 163.

Jackson, C. V. (1953). Visual Factors in Auditory Localization. *Quarterly Journal of Experimental Psychology, 5*, 52–65.

Jacobs, L.; Karpik, A.; Bozian, D. & Gothgen, S. (1981). Auditory-Visual Synaesthesia: Sound Induced Photisms. *Archives of Neurology, 38*, 211–216.

Kluver, Heinrich (1966). *Mescal and Mechanisms of Hallucinations*. Chicago: University Press.

Lewkowicz, David J. & Turkewitz, Gerald (1980). Cross-Modal Equivalence in Early Infancy: Auditory-Visual Intensity Matching. *Developmental Psychology, 6*, 597– 607.

Lipscomb, Scott David (1995). *Cognition of Musical and Visual Accent Structure Alignment in Film and Animation*. Unpublished Diss. University of California, Los Angeloes, CA.

Marks, Lawrence E. (1978). *The Unity of the Senses: Interrelations among the Modalities*. New York: Academic Press.

Mateeff, Stefan; Hohnsbein, Joachim & Noack, Thomas (1985). Dynamic Visual Capture: Apparent Auditory Motion Induced by Moving Visuals Target. *Perception, 14*, 721–727.

Meredith, M. Alex & Stein, Barry E. (1986). Visual, Auditory, and Somatosensory Convergence of Cells in Superior Colliculus Results in Multisensory Integration. *Journal of Neurophysiology, 56*, 640–662.

Myers, Charles S. (1915). Two Cases of Synaesthesia. *British Journal of Psychology, 7*, 112–117.

Pourtois, Gilles & Gelder, Beatrice de (2002). Semantic Factors influence Multisensory Pairing: A Transcranial Magnetic Stimulation Study. *Cognitive Neuroscience and Neuropsychology, 13 (12)*, 1567–1573.

Radeau, M. (1994). Auditory-Visual Spatial Interaction and Modularity. *Cahier de Psychologie Cognitive – Current Psychology of Cognition, 13*, 3–51.

Shams, Landan; Kamitani, Yukiyasu & Shimojo, Shinsuke (2002). Visual Illusion Induced by Sound. *Cognitive Brain Research, 14*, 147–152.

Simons, Daniel J. & Chabris, Christopher F. (1999). Gorillas in our Midst: Sustained: Inattentional Blindness for Dynamic Events. *Perception, 28*, 1059–1074.

Stein, Barry E. & Meredith, M. Alex (1993). *The Merging of the Senses*. Cambridge, Mass.: MIT Press.

Thurlow, Willard R. & Jack, C. E. (1973). Certain Determinants of the Ventriloquism Effect. *Perception & Motor Skills, 36*, 1171–1184.