Audiovisual Perception in the Context of Creating Art

Helga de la Motte-Haber

Audiovisual perception is a mysterious process. It enables us to gain understanding about objects and events as well as about ourselves. However, there is seldom awareness that the way the senses function determines the section of what is experienced as one's "environment"[1] within a surrounding area, which is not the same as that of other living beings. The process of perception as such is accessible to inner consciousness but not as a clear notion. Franz Brentano elucidated this in his work of 1874 using the example of hearing: "We can observe the sounds we hear, but we cannot observe our hearing of the sounds for the hearing itself is only apprehended concomitantly in the hearing of sounds."[2] In today's terminology we would say the process of perception as such can hardly be separated from the associated processes of information processing, from experiencing, perceiving, and understanding. To unravel this information processing is the goal of cognitive science, psychology, and neurology; today it also plays a role in the development of artificial intelligence (AI). Instead of giving this introduction, perhaps I should have just said that what I am speaking about is highly complex and will only be rudimentary. In addition, the attempt to parallel scientific findings with art-making risks the reproach that one will hardly find a one-to-one relationship because the former has to set conditions with few landmarks in order to substantiate effects, whereas the latter's objective is a wide aesthetic impact. Nevertheless, in both are expressed intermodal experience and mutual influences between sense organs.

I. Cross-modal Equivalence

Intermodal Qualities

Intermodal qualities adhere to all unimodal perceptions. One experiences a sound as clear or a colour as garish. Full, empty, quiet, nervous, round, thick, dense, porous, and so on are the kind of characteristics that feature in different sensory impressions. This means that their material basis cannot be clearly specified. In his widely read work of many editions *Entwicklungspsychologie* of 1926,[3] Heinz Werner attributes these non-objective sensations to an expressive effect or physiognomic perception. He applied these thoughts to observations of *Sprachphysiognomik,* the physiognomic properties of language, to begin with in an essay of 1928, and then in a book published in 1932. He refers to "intersensory qualities", which "cannot be assigned to the optical, the acoustic, or the tactile by their

nature".[4] Werner assumed that expressive perception was primitive, or rather that children already had a developed understanding of physiognomic qualities.[5] After his emigration to the USA and the translation into English of his *Entwicklungspsychologie*, the term "tertiary qualities" appears and with it an alignment with the terminology of John Locke, that is, with empirical thought; however, this was in contradiction to Werner's commitment to the epistemological foundations of Gestalt theory and holistic psychology. The question of which ideas Heinz Werner received from the music educator Gertrud Grunow is a research derivative. Until 1924, she was the only female Master teaching at the Bauhaus, and subsequently she worked with Heinz Werner in Hamburg. She designed a colour wheel equivalent to the twelve-tone technique. Even in 2001 Gernot Böhme's concept of atmosphere[6] refers to Heinz Werner although he criticises Werner's view as intersensory.

Synaesthesias, which included intermodal qualities, were a prominent topic in the 1920s and can certainly be seen as connected to contemporary developments in art. Georg Anschütz, synaesthesia researcher, organised a total of four Colour–Sound Congresses (1927 to 1936). Albert Wellek published a report on the first Congress in 1928.[7] He referred to Werner and spoke of universal primordial human correspondences (also primordial synaesthesias).[8] Wellek was one of the few writers who cited Helmuth Plessner's epistemological treatise *Die Einheit der Sinne*[9] published already in 1923 (supplemented in 1980 with the *Anthropologie der Sinne*). Plessner was involved in teaching the science subjects that were part of the Bauhaus curriculum, thus it is understandable that he emphasised the relations between ear and eye and engaged with the musicalisation of the visual arts. Plessner posited the existence of a *sensus communis* (common sense), which Aristotle had described in *De anima* as an inner capacity that tells us what all the senses have in common. Further, he named overlapping qualities — intensity, brightness, volume, density, and roughness — which were taken over in the more recent texts and expanded. Seldom, however, is there any impression of warmth, with which sounds and colours can be endowed and not only kinaesthetic reactions. Smells, too, like those that Katja Kölle employs in her installations, are seldom mentioned. Incidentally, it seems that not all intersensory qualities are known.

Colour and Sound

In the first half of the twentieth century it was above all the synaesthetic connection between colour and sound that played an important role. In the meantime this relationship has been elaborated in depth, and extensively documented by Karin von Maur's major exhibition *Vom Klang der Bilder* (The sound of pictures) (1985).[10] However, it must be noted that the same

applies to the exhibition as Plessner remarked of Wassily Kandinsky and the Futurists: they are still seeking the true music for the eye "within the confines of the panel painting".[11] This has now been replaced by other forms.

An impressive colour and sound installation was created in 1970 by composer Alvin Curran with the painter Paul Klerr for two connecting rooms of a gallery in Rome.[12] The rooms were threaded with many meters of coloured strings made of various materials (metal, wool, nylon, etc.) in a web of geometric patterns that obliterated the normal perspective. Viewers could modify the installation's (colour) structure by changing the position of a number of strings attached to sliding intersections. In the second room groups of chimes, tubes of hard and soft metal of different lengths, were suspended to produce sound. In the corners some strings converged in small metal boxes reminiscent of guitars. In the first room there were no chimes, only one waxed wool string, which when rubbed was made to sound. The passage-way between the rooms was fitted with a resonating wooden frame strung with steel strings like a piano. The installation was miked so that the sounds created by the visitors could be heard in two other rooms. Clearly, the installation was inspired by the then new idea of audience participation. Its title, *The Magic Carpet,* evokes a colourful Persian carpet with intricate patterns and designs in which their constant movement gives rise to a kaleidoscopic fabric of colours and sound. Here synaesthesias were not a consideration, but rather holistic perception.

The Tonal Space

Intermodal qualities are manifested in "seeing sound", which is reflected in the typeface of scores by the spatial position of bright/high and dark/deep sounding instruments, whereby at the same time it seems to hold that high tones in contrast to deep tones are not round but appear pointed.[13] The spatial representation of sound is largely independent of the position of the sound source. Whether emitted by a loudspeaker placed above or below does not change the height impression of the tones at all.[14] The inner space of the tonal order, to which Hermann von Helmholtz drew attention,[15] is structured vertically similar to the external space. Common to both is that the same movements can take place there; that is, the same thing can appear at different positions within the space. This concerns primarily the vertical, which is metrically precise irrespective of pitch. The horizontal is, like vision, dependent on the position of the observer/listener. In addition, sounds possess object qualities, which Carl Stumpf[16] described as discrete dimensions besides the pitch impression of the tones as full-empty and dull-sharp. These qualities were confirmed in experiments by Taffetta M. Elliott,

Liberty L. Hamilton, and Frédéric Theunissen in 2013.[17] Why these qualities are frequently assigned to tone colour is unclear. This multi-dimensional acoustic space can modify the architectonic space by expanding it.

In addition to such primordial synaesthesias, learned associations, or connotations, play a role in assigning visual impressions to acoustic ones. In the documentary recording of sounds on the seashore at dawn, imperceptibly edited by Luc Ferrari to 21 minutes (*Presque Rien No. 1. Le lever du jour au bord de la mer*, 1967), such connotations, the quiet barking of a dog or faraway voices, call up images of the breaking day, and together with the echoing effects, of the entire space. Intermodal qualities adhere to the sounds themselves with their different levels of volume and clarity. The acoustic space with its echoing effects that Ferrari opens is entwined by the listener with impressions that operate visually.

The Tonal Space in the External Space

The sound sculpture *Conloninpurple* by Gerhard Trimpin of 1998[18] is dedicated to the meshing of inner and external tonal space, and is playable as an instrument in equal temperament. It consists of various spatial segments of closed and open wooden tubes of different lengths with funnel-shaped metal resonators, which are struck by electromagnetically controlled hammers. Pitches are arranged vertically in octaves and minor seconds (the latter ensure that an unending Shepard scale does not result), and with tubes in minor thirds a diverse overtone spectrum unfolds horizontally. Playing the sonically segmented tonal space in the real space effects its kaleidoscopic reshaping.

In his *Music for Passageways* of 1985[19] Robin Minard merged the vertical with the horizontal. Also tuned to equal temperament were the 32 tubes of different lengths with integrated loudspeakers for taped sound set up in a passageway. In this highly complex installation, starting from the four sides of the room the length of the tubes increased, crossed by the middle pitches in a curved line. This created a field of different pitches that might remind one of walking through a playing orchestra where the high and low instruments are approximately arranged differently in the real space.

Cross-modal equivalence, which allows listeners to experience sound spatially and sculpturally in real space, is a prerequisite of Bill Fontana's sound sculptures. In his first work *Kirribilli Wharf* (1976), made in the suburb of Sydney that lies on a promontory in the harbour, Fontana made an eight-channel field recording of wave action beneath the pier, directly above the source of the sounds. He soon changed to making Relocations — placing an

ambient sound source within a new context deemed appropriate in order to intensify the subtle interplay between the source of the sound and a life situation. In the years 1972–1976 Fontana transferred resonances from ambient urban sounds in glass vessels to the interiors of buildings. This intense engagement with resonances still imbues his work *Harmonic Time Travel* (2019) realised in Bonn, Germany.[20]

Artists' interest in inner and external tonal space was not an entirely new phenomenon in the twentieth century. Ludwig van Beethoven, for example, had challenged common notions of musical space in order to see over the "starry firmament" (Symphony No. 9) and he made the lowest pitched instrument in the orchestra, the double bass, play a tremolo on its highest note: it was this unstable sound on which Beethoven based his orchestral setting.

Further interesting aspects about modification of the expression of the external by the inner musical space can be discovered by an exact analysis of Gustav Mahler's Symphony No. 2. At one point (Resurrection) the offstage brass ensemble causes the musical space to disintegrate to such a degree that the concertgoer in the auditorium must be afraid of losing their seat. For the sake of brevity I will here refer readers to the analysis by Harald Hodeige (2004).[21] Anton Bruckner takes a completely different approach than Mahler by creating a special form of site specificity, because the general pauses in his symphonies actually embed the resonances of the real space in the musical structure. Here it should also be mentioned that intermodal qualities play a role in sonifications, namely, in a compositional context.[22] Concrete examples are the sonification of the shapes of mountain ranges. In the simplest case their elevation is translated into tone pitches and distances into durations. This parallelism provided the stimulus for the compositions *Gruppen für drei Orchester* (1955) by Karlheinz Stockhausen and Alvin Lucier's *Panorama* (1993).

II. The Different Capacities of Ear and Eye and Their Multimedia Interaction

As a rule the dominance of the eye is usually substantiated by citing the fundamental form of viewing, namely, of space. Reference is seldom made to the fact that one cannot view space as such, only objects around which light rays bend. When everything in the field of vision is the same colour and equally illuminated, visual perception breaks down; looking into this so-called ganzfeld one believes one is looking at infinity. This effect, experienced in outer space or in fog, was demonstrated in experiments by Wolfgang Metzger (1930).[23] In his *Ganzfeld Pieces* James Turrell translated it into art. In contrast to the eye, the ear is a regular analyser of space, for hearing encompasses all the surrounding space including one's own position in

it; that is, not only one visual channel, but the entire volume plus anything happening behind one's back. In evolutionary terms the ear is a warning organ because of its rapid reaction capability. We experience its ability to stimulate alertness on a daily basis. Before the findings of scientific research had been published on the subject, one could observe in the short American TV thrillers of the 1970s and 1980s how there would be a brief snatch of music before a sudden change of scene (back then music was only used sparingly in such TV productions because it was expensive). It was not necessary to interpose a shot of a street in order to transport the viewers from one scene of action to another. In the meantime we now know that in general a preceding sound of 100–400 milliseconds can arouse involuntary alertness to visual stimuli.[24] On the other hand, the reaction speed to acoustic stimuli cannot be changed by light signals.[25] The human ear can distinguish time differences of 1–2 milliseconds, whereas the eye needs a difference of 17 milliseconds. Possibly, the orchestral landscape would not have developed to the same extent if conductors had not been able to exploit this reaction capability for the synchronous playing of the musicians. Although one cannot hear as far as one can see into the distance, acoustic directional hearing is quite precise. Resonances signal the characteristics of objects in a space. However, a great many more indicators are available to the eye to identify objects. The ear, by contrast, can only draw conclusions from the volume and the frequency differences. When it is dark the eyes rely on the ears.

The Dominance of Vision

"For seeing, I'm born, for watching, employed".[26] Surprisingly this remark from Goethe's *Faust II* is made in a scene that plays in the dead of night. However, it is in agreement with the many investigators who assume the primacy of vision over hearing. Until very recently there were repeated attempts to prove the superiority of vision through the "ventriloquism effect". The perception of the source of sound is shifted onto something visible (within a variance of 30°). The best known example of this is the ventriloquist with his or her dummy. What is seen and heard is not simply represented but processed as information. In former times cinemas often had loudspeakers installed on the side walls of the auditorium, yet the film's protagonists always seemed to be speaking from the middle of the screen. The McGurk effect[27] also spawned many studies. It was demonstrated that heard consonants could be modified by seen movements of the lips ("ba" becomes "ga"). In the case of string instruments a seen bowed or plucked movement could at least unsettle the perception of the sound heard.[28] However, the McGurk effect is not so clear cut as is often maintained. It is

frequently suppressed that the conflict between the heard and seen information is also resolved by the fact that something new is heard; namely, “da”. This indicates that perception is fundamentally oriented toward constructions of meaning.

This was demonstrated in an impressive way by Hans Peter Kuhn. He used the ventriloquism effect in his 1994 sound and light installation *Jumps (Sprünge)* in Potsdam. In a former granary he set up a line about 50 meters long of 12 identical sections of large diameter steel tubing looking rather like bass drums; above each steel tube was a halogen lamp. Inside the tubes hung bass loudspeakers, which emitted very deep sounds, all overtones having been filtered out. Thus it was impossible to recognise from whence the sounds originated. Sounds were emitted at irregular intervals and distributed among the tubes at random. The view of the line of shining steel tubes lit up caught the eye and held it. The impression received was of sounds jumping along the line of speakers.

What You See Is What You Hear

Compared with the number of studies devoted to single sensory organs, the volume of research on audiovisual perception is still quite small, although it is growing all the time. As is to be expected, the topics are quite varied. However, some of the investigators are linked by a common element: refuting the hypothesis of the dominance of the eye. For even in everyday experience we can confirm that, irrespective of the position of the sound source, weak light coupled with a broadband auditory stimulus is perceived as brighter.[29] One article, only one page long, has the provocative title “What you see is what you hear”.[30] It was demonstrated in an experiment that a pulsating sound stimulus (60–100 msec) can make a uniform light flicker. In a row of white dots on a black background, sound can facilitate identification of objects in a search, while at the same time the dots appear brighter and of longer duration.[31] This latter “freezing effect” is, however, a quantitative effect; whereas the experiment with the beeps and flickering light indicates a proper quality change. In the meantime the flash effect has been investigated many times, and it appears that the sound source does not even have to be in the same place to influence the light.[32] Composers of film music have always known that they can intensify the experience of a chase in a film with fast rhythmic pulses. Listening to rhythmic sequences may even affect the speed of the eye’s saccadic movements (scanning).[33] Of the many cross-modal effects, which particularly benefit animation films, I shall only mention one more: the rather amusing bouncing effect. With two crossed diagonals of dots, one running from top left and the other from top right, Wolfgang Metzger demonstrated the Gestalt law of continuity.[34] However, when the

diagonals are presented in motion on a screen with a sound at the point where the lines cross, they do not continue to run on but instead change direction and remain on their side of the screen.[35]

Selective Attention

To process and interpret all the information we are bombarded with would lead to overload. Instead, attention is focused so that, for example, we can listen to one speaker in the midst of a babble of voices (cocktail party effect). However, this focusing gives rise to a form of "inattentional blindness", as shown in the well-known video by Daniel J. Simons and Christopher Chabris:[36] an observer has to count the ball exchanges in a game and does not notice when a gorilla runs across the frame. With respect to selective attention, hearing, as mentioned above, plays an important role in multimedia perception. It also partly explains the visual flash effect.[37] But hearing can also automatically stimulate attention for something unimportant and overlooked; that is, it can cancel the inattentional blindness effect. In a video where a pianist in a living room is playing all kinds of irrelevant stuff, a small cat sits on the side-lines and is only noticed when a small meow is heard.[38] Attention combined with multimedia effects plays a significant role in installations which accentuate space through sound.

One example of artistic accentuation of space is a work by sound artist Andreas Oldörp. In *!Horizont* (1998), presented in the Kunsthalle Krems, he used a row of different organ pipes, supplied with air through glass tubes and mounted along the side of a ramp, to direct attention along the wall. Surprisingly, at different times different notes would stand out against the overall sound,[39] as though they wanted to bring the standpoints of various pitches up close to the visitors wandering about.

Accentuating That Which Is Normally Overlooked through Listening: Soundwalks

The description of a place, as the authors of novels well know, is inadequate without auditory impressions. Nevertheless, it is quite a surprise that a sound protocol of a Viennese suburb was made in 1907 — a long time before the advent of and intermittent vogue for soundwalks — by the later director of the Burgtheater in Vienna, Baron Alfred von Berger.[40] The Japanese artist Akio Suzuki worked in a similar way to the Viennese Baron, but in his *Oto Date* project (1988 – ongoing)[41] he marks "listening points" he has chosen with a stencil of a

ring with a drawing of two ears inside it at locations in cities, which, although they are familiar and have been seen many times, give rise to the experience of seeing them for the first time.

Max Neuhaus was the first[42] to propose listening in public places. In 1966, *Listening* was the first soundwalk, which he took around Manhattan with friends who had the word “Listen” stamped on their hands. Neuhaus described the places and what people often failed to hear there. The latter presumably did not apply when walking past the “spectacularly massive rumbling” power station.[43] His subsequent “Lectures” he repurposed as soundwalks. Today, Neuhaus is often subsumed under the heading of acoustic ecology. However, his concern was, like one year later with his legendary *Drive-In Music*, that sound compositions must not be forced into a time frame. Ecologically oriented, by contrast, was the *Sound Scape Project* (1970) by R. Murray Schafer, and the *Sound Walks* of Hildegard Westerkamp were indebted to the usual desire at that time to foster sensitisation. Later, the Walkman, and today the apps on smartphones, became auxiliary ears. A new form of electroacoustic music developed, which ranged from sampling to using apps to locate the place of origin of a sound (cf. the ZKM).[44] Whether any of the pieces created in this way will dethrone Luc Ferrari’s *Presque Rien* remains to be seen.

“What you hear is more than what you see” — many of the very diverse installations by Christina Kubisch can be characterised like this. Through sounds the obscured traces of times gone by become apparent. Electromagnetic fields of cables, which had once carried sounds are fixed to the wall or wound around trees, whisper of their secret life that becomes audible when wearing special, sensitive wireless headphones developed by Kubisch. The sounds of flowing electric current which is invisible (e.g., in ATMs) and often from subterranean municipal facilities turn the participants in her *Electrical Walks* into seismometers. They then hear what they don’t see.

The Problem with Vertical Perception: The Ambiguous State of Research

Make a pile of 12 one cent coins,[45] measure the height of the stack, and compare it with the diameter of a one cent coin. The pile seems higher but it actually measures the same as the diameter of a one cent coin.

This visual illusion of perception, which was first described in 1851 by the physician Adolf Eugen Fick, is due to overestimating the vertical,[46] which is sometimes used by artists to give the impression of taller sculptures. It depends on integration in the field of vision and thus can be modified by the external framing. Overestimation of the vertical is sometimes

explained by learning effects, which suggests that blind people do not possess it. This was indeed demonstrated in an experiment with people blind from birth, who were tasked with estimating an acoustically implemented visual vertical.[47] However, the test subjects' estimates were based on acoustic conditions — was this a flaw in the experiment? With regard to anthropological purport of optical overestimation of elevation, environmental psychologist Rainer Guski surmises that it results from the need to remain safe in case an object falls.[48] This explanation is hardly compatible with the assumption of learning effects.

Acoustics can influence the visual vertical. For example, visual perception of the apparent verticality of glowsticks in the dark was shifted away from the side where sound was coming from.[49] Unlike the eye, the ear is not subject to the elevation illusion with regard to ambient sounds, but sometimes when headphones are worn.[50] The impression of visual verticals, which are intended to sound quasi-natural, is difficult to construct artificially. Although 3D systems for spatial auditory illusions have been developed, in his recent book on electroacoustic music, Curtis Roads[51] says they are less suitable for composing the vertical dimension of the listening space than the traditional overhead speakers. For his epoch-making *Gesang der Jünglinge* ("Song of the Youths") of 1956 with its fusion of the human voice and electronic sounds, Karlheinz Stockhausen[52] removed the overhead speakers, still included in 1955, and distributed the sound all around the listeners. He did not state any reasons for doing this. In the meantime there are again many compositions that utilise sound from above. Notable examples are *Kugelauditorium* (1970) by Karlheinz Stockhausen and Luigi Nono's *Prometeo* (as of 1984). Here it is particularly spatial surround sound that plays a role. Perception of the vertical in sound installations is generally not specially structured, but left to the propagation of sound and its resonances in the real space. For his *Music for Environmental Sound Diffusion* (1984) Robin Minard also used 10 ceiling loudspeakers, but to produce a homogenous sound impression in the room.

Perhaps it took a master builder to make verticals impressive. Bernhard Leitner, a qualified architect but who uses sound as building material, is one of the few artists who has created verticals in several works. In the 1970s, the listener for Leitner was a small column in the midst of ascending and descending sound. In this, the head still had to be capable of making its small, normal movements, for without differences in intensity and the time lag between the two ears, a kind of "blind spot" for what is above develops. In the intervening years Leitner has built veritable sound cupolas, and in 2005 even as an acoustic reconstruction of a semi-ruined church in Berlin. The sound of a trombone, as "heard" in classic depictions of angels, rose up, swelled above with prolonged reverberations (48 sec), sank down, and rose again.

Also impressive are Leitner's parabolic condensations with which he constructs columns of sound in rooms.

Short Summary

Practically all art forms interrogate the relationship of fictionality and reality. Contemporary art generates an awareness of our perception's construction of meaning. This, art has in common with science; science, however, searches for explanations of why we see things in the way they appear to us. Art, by contrast, points out the possibilities of interpretations that go beyond such constructions of meaning and their potential for opening up new experiences.

Translated from the German by Gloria Custance

Notes

[1] The term "Umwelt", "environment", which had not been in extensive use before, was shaped in its modern meaning by German biologist Jakob Johann von Uexküll at the beginning of the twentieth century.
[2] Franz Brentano, *Psychologie vom empirischen Standpunkt*, vol. 2 (Leipzig: Duncker & Humblot, 1874) 159; English translation: *Psychology from an Empirical Standpoint* (London: Routledge, 1995) 134.
[3] Heinz Werner, *Einführung in die Entwicklungspsychologie* (Leipzig: Barth, 1926) 45ff; English translation: *Comparative Psychology of Mental Development*, trans. E.B. Garside (New York: Harper & Brothers, 1940).
[4] Heinz Werner, *Sprachphysiognomik* (Leipzig: Barth, 1932) 6.
[5] Ibid, p. 9.
[6] Gernot Böhme, *Aisthetik: Vorlesungen über Ästhetik und allgemeine Wahrnehmungslehre* (Munich: Fink, 2001) 90.
[7] Albert Wellek, "Farbe-Ton-Forschung und ihr erster Kongress", *Zeitschrift für Musikwissenschaft* 9 (1928): 576–583.
[8] Albert Wellek, "Das Doppelempfinden in der Geistesgeschichte", *Zeitschrift für Ästhetik und Allgemeine Kunstwissenschaft* 23 (1929): 14–42.
[9] Helmuth Plessner, "Die Einheit der Sinne, ergänzt um *Anthropologie der Sinne*", in *Gesammelte Schriften*, vol. 3, eds. Günter Dux, Odo Marquard, and Elisabeth Ströker (Frankfurt am Main: Suhrkamp, 1923; 1970/1980).
[10] Karin von Maur, *Vom Klang der Bilder: Die Musik in der Kunst des 20. Jahrhunderts* (Munich: Prestel, 1985).
[11] Plessner, 1923; 1970/1980, 253.
[12] For a detailed description of *The Magic Carpet* installation with images and sound, visit Alvin Curran's website: www.alvincurran.com
[13] Thus there is no difference to languages, which describe high tones as pointed (*aigu*, sharp) and deep tones as ponderous (*grave*, heavy).
[14] Carroll C. Pratt, "The spatial character of high and low tones", *Journal of Experimental Psychology* 13, 3 (1930): 278–285.
[15] Hermann von Helmholtz, *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik* (Hildesheim: Olms, 1863/1968) 597; English translation: *On the Sensations of Tone*

as a Physiological Basis for the Theory of Music, trans. Alexander J. Ellis (London: Longmans, Green, 1885).
[16] Carl Stumpf, *Tonpsychologie*, 2 vols. (Leipzig: Hirzel, 1883, 1890) 530ff; English translation: *Tone Psychology*, vol. I, ed. and trans. Robin D. Rollinger (London: Routledge, 2020).
[17] Taffetta M. Elliott, Liberty L. Hamilton, and Frédéric Theunissen, "Acoustic structures of the five perceptual dimensions in orchestral instrument tones", *Journal of the American Acoustical Society* 133, 1 (2013): 389–404. A further factor, tonal-noisy, relates less to spatiality and more to tone colour. The weak fifth factor in this study could be an artefact.
[18] Anne Focke, ed., *Trimpin: Contraptions for Art and Sound* (Seattle, WA: Marquand Books, 2011) 92.
[19] Robin Minard, "Music for Passageways", in *Robin Minard: Silent Music. Zwischen Klangkunst und Akustik Design*, ed. Bernd Schulz (Heidelberg: Kehrer, 1985) 88–89.
[20] www.bonnhoeren.de>klanginstallationen-in-bonn
[21] Harald Hodeige, "Komponierte Klangräume in den Sinfonien von Gustav Mahler", *Musikwissenschaft an der Technischen Universität Berlin*, vol. 5 (Berlin: Technische Universität, 2004).
[22] Julia H. Schröder, ed., *Sonifikation: Transfer ins Musikalische* (Berlin: Berliner Gesellschaft für Neuen Musik, 2017).
[23] Wolfgang Metzger, "Optische Untersuchungen am Ganzfeld II: Zur Phänomenologie des homogenen Ganzfelds", *Psychologische Forschung* 13 (1930): 6–29.
[24] Jon Driver and Charles Spence, "Crossmodal attention", *Current Opinion in Neurobiology* 8, 2 (1998): 245–252.
[25] Maurice Hershenson, "Reaction time as a measure of intersensory facilitation", *Journal of Experimental Psychology* 63, 3 (1962): 289–293.
[26] Wolfgang von Goethe, *Faust II*, Act 5, Scene 4; English translation https://www.poetryintranslation.com/PITBR/German/FaustIIActV.php#Act_V_Scene_IV
[27] Harry McGurk and John Macdonald, "Hearing lips and seeing voices", *Nature* 264 (1976): 764–748.
[28] Helena M. Sadãna and Lawrence D. Rosenblum, "Visual influence on auditory plucked and bowed instruments", *Perception and Psychophysics* 54, 3 (1993): 406–416.
[29] Barry E. Stein, Nancy London, Lee K. Wilkinson, and Donald D. Price, "Enhancement of perceived visual intensity by auditory stimuli: A psychophysical analysis", *Journal of Cognitive Neuroscience* 8, 6 (1996): 497–506.
[30] Ladan Shams, Yukiyasu Kamitani, and Shinsuke Shimojo, "What you see is what you hear", *Nature* 408 (2000): 788.
[31] Jean Vroomen and Beatrice Gelder, "Sound enhances visual perception: Cross-modal effects of auditory organization on vision", *Journal of Experimental Psychology. Human Perception and Performance* 26, 5 (2000): 1583–1590.
[32] Hamish Innes-Brown and David Crewther, "The impact of spatial incongruence on auditory-visual illusion", *PLoS One* 4, 7 (2009): e6450.https://doi.org/10.1371/journal.pone.0006450
[33] Jared E. Miller, Laura A. Carlson, and, J. Devin McAuley, "When what you hear influences when you see: Listening to an auditory rhythm influences temporal allocation of visual attention", *Psychological Science* 24, 1 (2013): 11–18.
[34] Wolfgang Metzger, "Beobachtungen über phänomenale Identität", *Psychologische Forschung* 19 (1934): 1–60.
[35] Robert Sekuler, Allison B. Sekuler, and Renee Lau, "Sound alters visual motion perception", *Nature* 385 (1997): 308.
[36] Daniel J. Simons and Christopher F. Chabris, "Gorillas in our midst: Sustained inattentional blindness for dynamic events", *Perception* 28, 9 (1999): 1059–1074.
[37] Jyoti M. Ramanathan, Antigona Martinez, and Steven Hillyard, "Effect of attention on early cortical processes associated with the sound-induced extra flash illusion", *Journal of Cognitive Neuroscience* 22, 8 (2009): 1714–1729.
[38] Daria Kvasova, Laia Garcia-Vernet, and Salvador Soto-Faraco, "Characteristic sounds facilitate object search in real-life scenes", *Frontiers in Psychology* (2019): 10, 2511, https://doi.org/10.3389/fpsyg.2019
[39] Diether de la Motte, "Neues Klangerleben mit Andreas Oldörp", *Positionen* 38 (1999): 56–57.
[40] Peter Payer, *Der Klang der Großstadt: Eine Geschichte des Hörens, Wien 1850–1914* (Vienna: Böhlau, 2018) 15.
[41] Suzuki created the first listening point in 1988 and titled it *Space in the Sun*.

[42] Max Neuhaus, https://doingit.fba.up.pt/en/max-neuhaus
[43] Afterwards, according to Neuhaus, they went back to his studio and he played the drums.
[44] For more information see the ZKM's platform of geolocalised sounds; project website http://mycity-mysounds.zkm.de
[45] Originally demonstrated by Wolfgang Metzger using one pfennig coins.
[46] For example, Gateway Arch National Park, St. Louis, MO, 1965.
[47] Laurent A. Renier, Raymond Bruyer, and Anne G. De Volder, "Vertical–horizontal illusion present for sighted but not early blind humans using auditory substitution of vision", *Perception and Psychophysics* 68, 4 (2006): 535–542.
[48] Rainer Guski, *Wahrnehmen: Ein Lehrbuch* (Stuttgart: Kohlhammer, 1996) 236.
[49] Floyd L. Ruch and Philip G. Zimbardo, *Psychology and Life* (Glenview, IL: Scott, Foresman & Co., 1971), p. 290.
[50] Dennis Folds, "The elevation illusion in virtual audio", *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50, 16 (2006): 1576–1579. In this study headphones have an effect on high tones; this is possibly a localisation effect of the headphones.
[51] Curtis Roads, *Composing Electronic Music: A New Aesthetic* (Oxford: Oxford University Press, 2015) 263.
[52] Karlheinz Stockhausen, "Gesang der Jünglinge", in *Texte zu eigenen Werken, zur Kunst Anderer*, 10 vols., vol. 2 (Cologne: DuMont, 1964) 49–70.