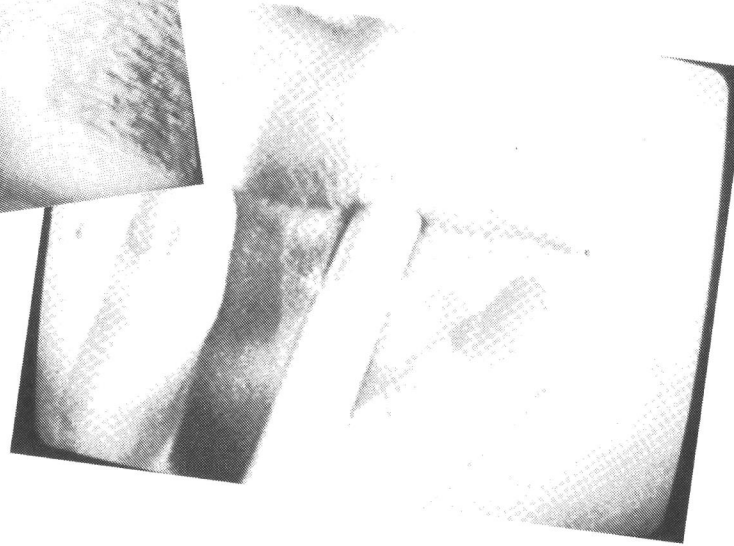
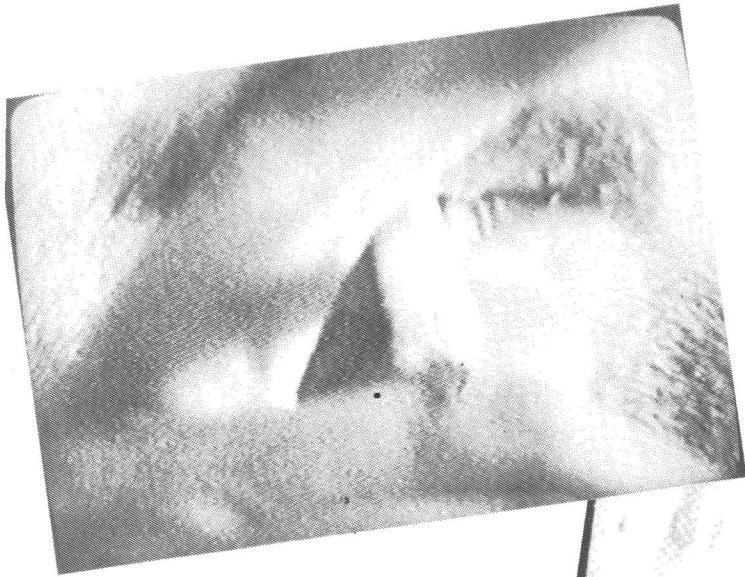


Simo Allitalo:

Sound and Image: Some Preliminary Considerations of an Audio-Visual Kuleshov Experiment



This study is based on an experimental video* that was made some years ago with Tuike Allitalo and Tapio Onnela. The video had mixed aims, we were supposed to learn how to use simple video equipment, to study sound and image relationships, to make an essay about the art of smoking cigarettes in the movies and to pay homage to Howard Hawks whose films I and Tuike were studying at that time.

The things I now want to discuss concern the sound/image or to be more specific music/image relationships of this video. The things which I will focus on in my paper are: what are the relations between the story space and the music space or what is the relationship between diegetic source music and traditional background music? Is it image alone that determines whether the music we have heard is diegetic or background

music? Is the common critical division between diegetic source music and extradiegetic background music too crude? If musical clues, independently of the images, inform our interpretations of the source of the music, what then are the musical mechanisms behind these interpretations?

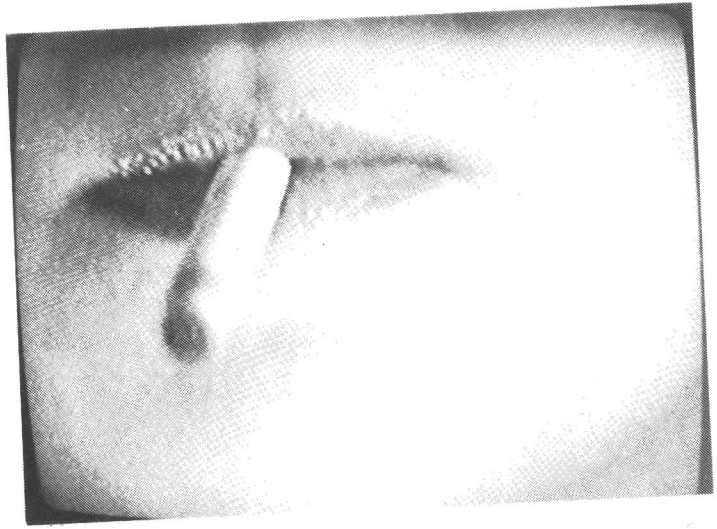
The following thoughts are based on my own perceptions and on a comments of a number of friends and neighbours. The results of our video experiment turned out to be a bit of a surprise. Especially the effects that music seemed to have on the meanings of the audio-visual whole.

The music was chosen haphazardly from my record shelf. The only principle was that the music should be as different as possible. The music used consists of fragments of:

1. Rodion Schedrin's Carmen-Suite based on Bizets Carmen.
2. Charlie Parkers live recording of Thelonius Monk's Round about midnight.
3. Edgar Varese's Ionisation.
4. An old Neapolitan song.

The joining of music to the image was done at random. No effort was made to synchronise the mixing or editing in any way. Actually, with the equipment avai-

* The video "Art of smoking" consists of four 2 minute extreme close-ups of a woman and a man smoking cigarettes. These were edited with the VHS video camera in sequence / woman / man / woman / man. The music was added afterwards. The images were shot in color.



lable, it would have been almost impossible.

The first question that seems to arise during the video, at least some people have reported that it was their first thought, is where does the music come from? Or on what level of narrative does the music belong? Does it emanate from the same space that the characters inhabit or does it belong to the extradiegetic level of the narrator? The same question could be formulated: Do the characters of the image hear the music or is it only available to us?

Reactions and interpretations of the spectators seemed to form a certain kind of a pattern. Although I must emphasize that the work was not done in strict empirical fashion so that the ideas and conclusions are only tentative.

The first excerpt (woman smoking & Carmen suite) was interpreted as an image with extradiegetic sound, in other words as normal background music. But some spectators said that the music gave them an impression that it was somehow connected with what the woman in the image saw. It was either understood as background music to the image of the woman seeing some dramatic, "operatic" things, or as the woman watching opera, ballet or some dramatic action accompanied by music from the diegetic space.

The second excerpt (man smoking & Round about midnight) was interpreted as an image with diegetic sound. Some people construed it as the man was sitting in a jazz club and listening to a band. The applause at the end of Parker's solo is evidence to this direction, but on the other hand the poor quality of the recording (the tape I used was second generation copy of a not so good LP-record) seems to indicate that the man is listening to a worn out cassette. Nobody noticed this possibility.

The third excerpt (woman smoking & Varese's Ionisation) was interpreted as an image with an extradiegetic sound. In addition, this example generated some evaluative remarks concerning the possible "genre" of the excerpt. It was considered as very arty-crafty piece indeed. Either the juxtaposition of the image and sound was considered somehow improper or it called forth some unpleasant memories of high-brow art. Some watchers reported that it reminded them of the pretentious way modern music was used in film during the late forties and fifties.

The fourth excerpt (man smoking & Neapolitan song) gave rise to some very interesting reactions. It was interpreted as Mind-Music. Either it was inner diegetic music that at least one person in the story spa-

ce, the man smoking, could hear, or it was metadiegetic background music that reflected the characters' aural thoughts.

The interesting thing in these reactions was that although there was no apparent evidence of sounds having a source within the story space some of them were construed as diegetic sounds. How come one interprets one piece of music as diegetic source music and another tune as background music when the image is almost exactly the same? The standard conception of film music is that, as long as there are no cues that music has a source within the story space of the film, it is to be understood as background music.

Now the question arises whether the standard conception is a solid one? If it is, one should be able to find visual or auditory cues that explain why we interpret sound/image combinations so differently in four video excerpts. If the signs or cues are visual they seem to be somewhat different from what is usually thought as an index of sounds source. In these images one does not see any radios, symphony orchestras etc. Other possibility is that cues are auditory. If so, are they auditory or specifically musical cues? By auditory cues I mean the noise of the medium, for example rattle or hiss of the recording, sound or echo of the location etc. Musical cues are based on specifically musical characteristics like rhythm, tonality, stylistic features. The third possibility is that the cues are based on cultural codes that underlie images, sounds and audiovisual complexes.

On the other hand, if the cues or signs of source location are musical it seems that the standard conception of film music is false or at least vague. But how could music be connected with the diegesis or story space even if there are no cues in the image of the source of the music being within the diegesis? One possible answer is that when people narrativise audiovisual complexes they are building possible worlds. In their world making practise spectators try build a coherent whole out of sounds and images. When doing this they rely on notions concerning the probability of certain sounds in specific environments.

And to be more specific about the problem that puzzles us, the question is not whether the music that accompanies the images is diegetic or extradiegetic but the hypotheses that spectators think are most probable. The next pictures may uncover the music having its source within the story space. But what interests us now is why spectators make hypotheses although they have no explicit evidence. Or maybe they have evidence that lies embedded in the music. Spectators seem to have some ideas about what kind of music could exist in this or that possible world.

II

Rick Altman writes in his article *Moving Lips: Cinema as Ventriloquist* (1):

"an individual who speaks will in all probability be the object of the camera's, and thus of the audience's gaze. In the political world, the right to free speech conveys a certain political power; in the narrative world, the right to speech invariably conveys narrational power for by convention it carries with it a secondary right, the right to appear in the image...., in general we may say that the actors gain the right to a place in the image by virtue of having previously obtained a spot on the soundtrack. I speak therefore i am seen."

Altman's ideas are very general indeed and in his article he is concerned mainly with the relationships between speech and image. But there are some inter-

esting points that I try to develop. In addition to the freedom of speech there are some other interesting legal/constitutional issues in narrational world. It seems that upon its arrest by the spectator an image has the right to remain silent, but everything it says may be used against it.

Now it puzzles us why the actors or people in the video image do not use their right to speak why do they remain silent, if they do not speak why are they to be seen? We start to question the audio-visual complex or should I say simplex. And because the image after its arrest does not remain silent we use everything it says against it. It seems that the actors' right to be seen in the image is not only implied by the fact that they are producing sounds but because they are receiving sounds. Although the persons we see smoking are not speaking we interpret them as acting, as doing something (other than smoking cigarettes; it might be that without the music they would simply be smoking cigarettes). We tend to narrativise what is happening because as Noel Carroll says: Narrative is ... the most pervasive and familiar means of explaining human action. (2)

It is interesting that the spectators' interpretations of the audio-visual complexes seem to indicate what the actors in the images were doing. The question "where does the music come from?" could be transformed to a question concerning the nature and object of the action only implicitly visible in the image. I think that we could condense these questions as follows:

1. excerpt What does she SEE?
2. excerpt What does he HEAR?
3. excerpt What does she ?????
4. excerpt What does he THINK?

As we remember the spectators interpreted the third excerpt as an image with background music emanating from outside of the story space. So it did not raise the question "where does the music come from?" and it seems logical that we can not reconstruct a question concerning the object of the action. This may be the reason why some spectators felt that the audiovisual sign was somewhat narratively empty or artificial. But this example was the only one that gave rise to value judgements concerning the style and possible "genre" of the excerpt. I think that the both above mentioned things are interconnected and depend upon certain cultural codes and our musical competence.

III

Simon Frith has said that "To develop the theory of film music we need in Antoine Hennions words, 'not so much a sociology of music as a musicology of society'"(3). I think Hennion ment that it is not enough to know how people and societies use music, but we need also to know how music uses us. I dont know wheter Hennion has developed his idea any further from this slogan but I think that Gino Stefani's book *La Competenza Musicale* (4) contains some ideas that could take us a little further towards the theory of film music.

According to Stefani, musical competence means ability to recognize or build structures, correlations etc. between music and surrounding culture. Musical competence functions on different levels, and various cultures and sub-cultures give more weight to some levels and disregard others. Stefani's model of musical competence contains the following levels.

General codes:

The basic perceptual and logical schemes by which we

perceive and interpret everything we encounter in the world, including our hearing experiences. "This could propably be called the 'anthropological' level were the term not so full of implications." (G. Stefani)

Social Practices:

This level refers to different cultural institutions like language, dressing, industrial work, spectacles etc. and among them the musical institutions like concerts, music critique etc. The history of music is full of examples which manifest the influence of social practices in music: march, hymn, lullaby, serenade, entrada, prelude, etc.

Musical Techniques:

This level refers to musical practices like instruments, scales, compositional forms, music theory. At this level music emerges as autonomous art.

Styles:

This level refers to the style of the period, school or an individual composer, taken together or separately. It designates the specific ways the musical techniques, social practices and general codes are put to use.

Works:

The level of unique musical works/compositions. The minimal competency at this level is the ability to recognise musical works.

According to Stefani, these levels are not hierarchic but they function in coexistense. But different musical cultures and aesthetics tend to give more value to certain levels and appreciate others less, e.g. in western musical culture different musics and various audiences prefer/use musical competences that differ in emphasis given to the levels of reception.

The interpretations our subjects made of sound/image montage may differ because the music used called forth different levels of musical competence. It is possible that the second excerpt (Charlie Parker & man smoking) was interpreted according to the social practices that we connect to jazz music: jam sessions, jazz clubs etc. It is customary or even a cliché to connect jazz with small and smoky milieus.

I think that all four examples could be analysed to the detail using Stefani's model as a starting point.

One more thing: as I said we use different levels of musical competence when trying to make sense of various musical genres, lets say opera and jazz. But listeners may also use differing levels of competence when trying to make sense of the same music.

German musicologists have criticised the way Stanley Kubrick uses György Ligeti's music in 2001 Space Odyssey. They say that nobody who has heard Ligeti's Lontano in cinema can listen it as a "pure" music anymore. This may be due to the fact that they interpret the music in the film only at the level of the work (cf. Stefani). Which is a mistake, at least in my opinion.

I think the way Kubrick uses music in 2001 and also in *Shining* tends to emphasise the level of general codes. And on this level he seems to operate especially with the music/noise opposition.

Because the lack of space my presentation of Gino Stefani's theory of musical competence has been quite fragmentary and I hope that Stefani will not be accused because of my simplifications. Interesting feature in Stefani's theory is that it seems to offer us at least a glimpse to the musicology of society which according Simon Frith is essential if we want to develop the theory of film music.

IV

There are some problems that undermine the assumptions made above. First: the vertical montage of sound and image does not completely fulfill the demands of the original Kuleshov experiment. The image is not exactly the same in all excerpts. One could maybe discern some minimal acting that would distinguish the takes without music along the same lines as the takes with music. The spectators might narrativise the images without music in some way isomorphic to the narrativisations explained above. If this were to be true then the image alone would determine the spectator's interpretations.

Another problem is the length of the takes. Each of the four takes is approximately 2 minutes long. Close-ups of such length are extremely rare in film and TV. So it is difficult to say whether it is possible to make any generalisations on the basis of this material.

There are some other things concerning the sound that were not taken into account during the preparation of the video. The volume levels and the dynamics of the sound were chosen randomly and no attention was paid to the reverberation, sound quality of the recordings etc. All these and many other parameters of sound could be varied and thus we could produce great number of minimally differing montages of sound and image out of the same basic material.

I understand that the notions I have presented to you are problematic because they are not based on experimental evidence. I am not saying that the narrativisations of the sound/image montages on which my reflections are based are the right ones or even the only possible ones. I became interested in this matter only because the spectator reports of what they saw and heard seemed to form a pattern. What interests me is whether it is possible to develop this audio-visual "Kuleshov experiment" into a experimental device which allows empirical testing of the ways the spectators perceive and interpret basic audio-visual gestures. This may prove to be a futile enterprise but I think that at least it can be used as a heuristic tool when studying the relationships of sound and image.

References

1. Rick Altman: *Moving Lips: Cinema as Ventriloquism*. Yale French Studies No 60, 1980. Special No. "Cinema/Sound".
2. Noel Carroll: *The Power of Movies*. Daedalus Fall 1985. Special No "The Moving Image".
3. Simon Frith: *Mood Music*. Screen 25 (3) 1984. Special No. "On the Soundtrack".
4. Gino Stefani: *La Competenza Musicale*. Universitaria Editrice, 1982 Bologna. All references are to the Finnish translation of Stefani's article which has appeared in a book: G. Stefani: *Musiikillinen kompetenssi. Miten ymmärrämme ja tuotamme musiikkia*. Jyväskylän yliopiston musiikkitieteen laitoksen julkaisusarja A: tutkielmia ja raportteja, no 3. 1985 Jyväskylä. All Translations are mine.

Stefani's article has also been published in Germany: G. Stefani: *Eine Theorie der musikalischen Kompetenz*. In Hans Werner Henze (Hrsg): *Die Zeichen. Neue Aspekte der musikalischen Ästhetik II*. Fischer Taschenbuch Verlag. 1981 Frankfurt am Main. After the Turku seminar I found out that an English version of the article: G. Stefani: *A theory of musical competence has appeared in Semiotica 66 (1/3) 1987. Special Issue: "SEMIOTICS OF MUSIC"*. Guest editor: Eero Tarasti.