ONE MICROPHONE AT THE MUSICIAN´S EAR: SPECIAL RECORDINGS OF THE MALAWIAN FRICTION BOW

Manfred Bartmann
University of Salzburg, Austria

ABSTRACT

Background. In 1992 Auhagen and Gätjen proposed a new method for recording the sounds of musical instruments by placing a microphone in contact with, or in close proximity to the musician´s ear. The ear was considered to be the most reliable place for gathering information concerning preferred sound parameters. It seems that such an approach is of great value in dealing with foreign musics. Ethno-musicological field recordings are usually played back to the player or singer, but seldom there is an opportunity for different microphone positions to be evaluated by the players themselves.

Aims.

- To allow analysis to be guided by the player´s own idea of sound.
- To establish a more reliable recording standard for comparative purposes.

Methods. Special 2-channel recordings of the playing of Malawian ethnologist Moya Aliya Malamusi, spectrographic analysis, special filterings derived from a psycho-acoustical overmasking paradigm, special videos for the evaluation of the element of movement involved.

Results. Recordings from the "at the ear" position tend to be preferred by the musicians. Spectrographic Analysis reveals that formants appear more pronounced in these recordings. High frequency as well as noise components are diminished. These findings were verified by using a model of simultaneous masking (Eckel 1989). By means of an overmasking function it was possible to convert the signal recorded at the musician´s ear into the reference signal which was recorded in front of the musician.

Conclusions. This recording method is of particular significance to types of music that are chiefly relevant to the musician himself - such as mouth bow music. Beyond these genres the method is more generally of use in shedding light on varying conceptions of sound. It offers the chance for an evaluation of intended figure/ground relations from a psycho-acoustical as well as from an intra-cultural point of view.

1. INTRODUCTION

The nyakazeze is a mouth-resonated friction bow which "is found among Sena and Nyungwe people in the area of the lower Shire and Zambezi rivers in Malawi and Mozambique. While the string which is shaped like a tape passes through between the lips, the musician´s mouth serves as a variable resonator. By changing the size of the cavity of his mouth, ... (the player) reinforces different sounds. In his right hand he holds a stick which is used for rubbing the stave of the bow. The left hand has also something to do. Its work is to hold the stave, and at the same time use two fingers to stop the vibration of the tape ... at certain points." (Malamusi 1999, Track No. 23, booklet p. 26).

Malamusi´s description makes it clear that this music is mainly intended for the musician himself. In order to find out whether this interpretation correlates with sound structures recorded at different positions, one microphone was positioned at a distance of 1 meter in front of the musician, while a second was placed close to his ear, i.e. at the left arm of an empty frame of glasses. In course of conversation Gätjen learned that recordings at the left ear position were preferred by the majority of the musicians he worked with. This correlates with the playing technique, in that the right ear position would provide mainly scratching noises. According to Gätjen, this recording procedure also helps to minimize the effects of sound reflections as far as normal rooms are concerned (Gätjen 1996). The method consequently considers the musician himself as a part of his instrument. Because it is him who continuously tries to optimize its timbre while playing it, his ears are considered the most reliable places where to learn about the preferred parameters of sound.

2. SPECTROGRAPHIC ANALYSIS

2.1. The Formant Paradigm

The contrast between the ear and the front position became most audible if I switched the two channels of the recording. On the stimulus level the differences mainly occur in the frequency band between 0 and 3 kHz (Figure 1). Higher ranges of the spectrum (not to be seen in Figure 1) are filled up with stave rubbing noises. The overtone contours between 0 and 1800 Hz carry the intended melodic pattern which is repeated many times in front of a fast elementary pulse provided by the regular scratching movement of Malamusi´s left hand. When recorded in the front position, these noisy components concentrate in a frequency band between 2 and 3 kHz (Figure 1, below), which comes out weaker at the ear position. There the mouth resonated contour even appears in a lower frequency band (Figure 1, above, 200-400 Hz).

Being traces of the areas of resonance provided by certain positions of the mouth cavity, these findings support the theory that these so-called formants not only dominate the timbre of vocal sounds but are also significant in instrument timbres (Schumann 1929, Fricke 1994). Needless to say, that in case of the nyakazeze continuous changes of timbre actually make up the music itself. The fundamentals of the string are simply means to an end. The fundamental pitches musically irrelevant, formants being pitch independent. Consequently, playing a mouth bow could be described as "tracing formants".

ISBN 3-931852-67-9
ISSN 1617-6847

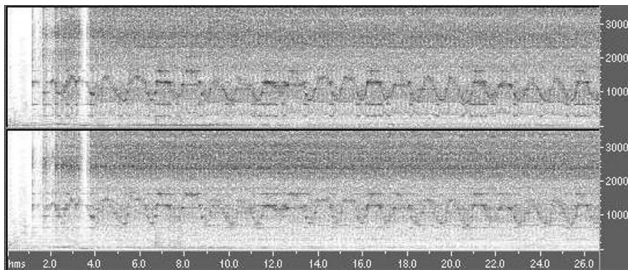

Figure 1: Malamusi playing the friction bow nyakazeze at Salzburg University (Institute for Musicology, seminar room, 14.01.2000). Spectrograms (16 kHz, 16-bit, 1024 Bands, logarithmic energy plot with a range of 80 dB) of the signals recorded at his ear **(above)** and at a distance of one meter in front of him **(below).** Most of the differences come out between 0 and 3000 Hz. The musically relevant overtone contour appears more pronounced at the ear position.

2.2. The Masking And Overmasking Paradigm

The question remains, whether the above mentioned findings may be confirmed by exploiting the masking properties of the human ear. The nyakazeze recordings do indeed suggest a verification of the formant paradigm by the masking and overmasking paradigm.

In 1989 Gerhard Eckel implemented a model of simultaneous masking for the analysis of musical signals (Eckel 1989). ”He used a modification of the spreading function which has been developed by Schroeder et al. (1979) in order to compute masking effects.” (Deutsch/Eckel/Noll 1992:2). I used the software package S_TOOLS© (PRG Deutsch/Noll 1992) in order to study the perceptual aspects of overmasking the nyakazeze signal which was recorded at the ear position. Based on the phase-vocoder (Dolson 1986), this S_TOOLS© filter not only makes it possible to decompose any sound signal but also to make its constituent parts audible again.

I reduced the ear-signal systematically by overmasking it to the most prominent spectral amplitudes, producing a series of filterings by simply increasing the offset level of the spreading function (Deutsch/Eckel/Noll 1992:2). I went as far as to remove on an average 92% of all spectral amplitudes (Figure 2, above). As a result, the full frequency range of the intended overtone contours still remained intact. Although the sound quality had dropped down dramatically, the signal still transferred the intended melodic pattern.

This comes as no surprise, as the most prominent spectral amplitudes can be assumed to be perceptually most relevant. In Western music leading voices can be extracted and separated from the sound of the accompanying orchestra (Deutsch/Eckel/Noll 1992). From this cultural perspective, overmasking to the most prominent peaks turns out to be a content-related procedure. It is confirmed by the figure-background organization common in Western Music.

Apparently this is only partly true for non-Western musics. In most parts of Southern Africa pitch and timbre are not necessarily understood as distinctive features of sound. As a rule, bow music is understood in a more holistic way, i.e. the organic relation between any resultant patterns and the whole is emphasized to a greater extent. In Botswana the rendition of similarly reduced bow music stimuli were regarded either as unintelligible or simply as “nothing”, i.e. as meaningless (own field experiments in 1999). As far as a listener has been encultured by the same or similar characteristics of music, i.e. as far as “African ears” are concerned, the reduced signal, although still carrying the complete mouth resonated contour, is in danger of no longer being considered as bow music.

These important effects and influences of musical enculturation remind us of the fact that Western speech becomes unintelligible if the signal is reduced to the 8% FFT-points with the highest amplitudes. “Overmasking destroys the well defined shape of the speech formants which obviously carry most of the vowel information” (Deutsch/Eckel/Noll 1992:4). However, the lacking information should still be found in the difference signal (Figure 2, below), which can be made audible as well (“analysis by synthesis”).

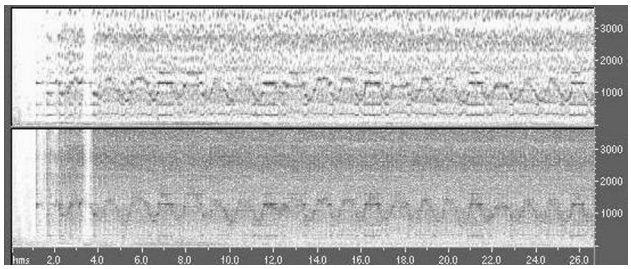

Figure 2: Decomposition of the signal shown in Figure 1 (above) by means of a perceptual based spectral peak picking method, derived from a psychoacoustical overmasking paradigm. Waveform-summation of these two items results in the signal shown in Figure 1 (above). **Above:** On an average, 92% of all spectral amplitudes are removed. Only the 8% FFT-points with the highest amplitude values remain present. **Below:** Waveform subtraction of the highest amplitude values from the original signal. This so-called difference signal provides almost the same result as the signal recorded in front of the musician (Figure 1, below).

3. ANALYSIS BY SYNTHESIS: CONVERTING THE EAR-RECORDING INTO THE FRONTAL RECORDING

Deutsch, Eckel and Noll (1992) recommend a discussion of overmasking in terms of figure-background discrimination. This corresponds with the fact that the most prominent peaks represent the intended figure. Subsequently the difference signal should contain most of the background information (Figure 2). Figure 3 suggests that this signal is perceptually equal to the one recorded in front of the musician.

ISBN 3-931852-67-9
ISSN 1617-6847

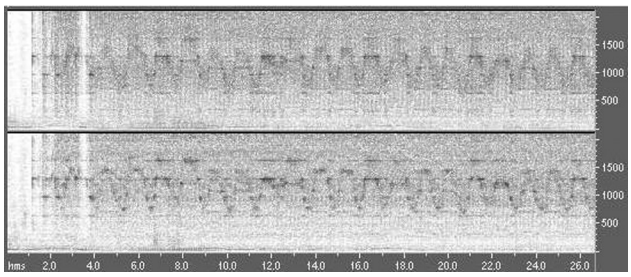

Figure 3: As a result of the filtering (overmasking) the less prominent spectral amplitudes make up a difference signal. It contains most of the background information of the signal recorded at the ear position (above). This constitutient part of the signal recorded at the ear is very similar to the signal recorded in front of the musician (below).

When switching the channels shown in Figure 3, the audible difference is mostly due to frequency components higher than 2 kHz. There is hardly any difference to be heard if both channels are reduced to a frequency band from 0 to 1800 Hz. This is the area which is of crucial relevance to the musician´s intention. There is a conspicuous coincidence between 0 and 500 Hz where in both channels no prominent overtone contours occur. This is the most distinguishing difference between the two recordings (Figure 1).

From a psychoacoustical point of view this recording method is useful in the clarification of some prominent aspects of figure-ground discrimination. The two recordings, as well as the filtered signals neither produce nor suggest any figure-ground inversion. In music perception, figure-ground discrimination apparently does not follow the principle of "exclusive allocation", which "says that a sensory element should not be used in more than one description at a time" (Bregmann 1990:12). Even the extreme overmasking procedure (Figure 2) suggests that there are transitional stages between what is fore- and what is background. Duplex perception is more likely than exclusion. This is because "sound is transparent", as Bregman puts it: "A sound in the foreground does not occlude a sound in the background the same way as a visual object in the foreground occludes our view of objects behind it. This means that a region of the spectrum that is localized as being at a certain spatial position can actually bear the mixed effects of two different acoustic sources." (1990:619).

As far as the playing of the nyakazeze is concerned, Gätjen´s (1966) finding that formants appear more pronounced at the ear position has proved true. This can also be verified by using the Viennese model for simultaneous masking. It is planned to pursue this sort of double-tracked investigation on other musics from Southern Africa.

4. OUTLOOK

Experimental research on aesthetic perception and response as well as on perception and psycho-physics often neglects musical concepts from non-Western cultures. This article is intended to shed a light on some new horizons. Having cultivated special devices or stratagies for acoustical concealment or deceit, quite a few African music traditions already helped to produce new insights concerning the figure-ground discrimination. In Southern Africa even ambiguities may occur (Bartmann 2002).

(1) In collaboration with Regina Sohneg (Institute for Musicology, Salzburg University, Austria) and Moya Aliya Malamusi (Oral Literature Research Programme, Chileka, Malawi).

5. REFERENCES

1. Auhagen, W. and Gätjen, B. (1992). Ein neuer Weg zur Aufnahme von Instrumentalklängen. *Instrumentenbau-Zeitschrift. Musik International 46. Heft 5,* (pp. 51-55).

2. Bartmann, M. (2002). Auditory Ambiguity in Xhosa Music: Interactions between Perceptual Complexity, Emic Concepts and Digital Storage. In Berlin, G. and Simon, A. (eds), *Music Archiving in the World* (pp. 269-277). Berlin: VWB (http://www.sbg.ac.at/mus/ muwi/institut/personal/bartmann_xhosa.htm; 05.04.2003).

3. Bregman, A. S. (1990) *Auditory Scene Analysis. The Perceptual Organization of Sound.* Cambridge: MIT Press.

4. Deutsch, W. A.; Eckel, G.; Noll, A. (1992). The Perception of Audio Signals Reduced by Overmasking to the Most Prominent Spectral Amplitudes (Peaks). *AES Preprint No. 3331, presented at the 92nd AES Convention 1992 March 24-27, Vienna.* (8 pages). New York.

5. PRG Deutsch, A./Noll, A. (1992). S_TOOLS© Version 5. Austrian Academy of Sciences. Acoustics Research Institute. Vienna (http://www.kfs.oeaw.ac.at/fsf/ dsp/index.html; 05.04.2003).

6. Dolson, M. (1986). The Phase Vocoder. A Tutorial. In *Computer Music Journal 10:4,* (pp. 14-27).

7. Eckel, G. (1989). *Ein Modell der Mehrfachverdeckung für die Analyse musikalischer Schallsignale.* Phil. Diss. Vienna.

8. Fricke, J. P. (1994). Ein Plädoyer für die Formanten. In Fleischhauer, G. et al. (eds) *Flöten, Oboen und Fagotte des 17. und 18. Jh.. 12. Symposium zu Fragen des Instrumentenbaus 1991 (= Beiheft 14 zu den Studien der Aufführungspraxis und Interpretation der Musik des 18. Jahrhunderts),* (pp. 66-77). Stiftung Kloster Michaelstein: Blankenburg, Michaelstein.

9. Gätjen, B. (1996). Auf das Ohr des Instrumentalisten bezogene Klangaufnahmen – ein Weg zur Erzeugung besserer Instrumentalsamples? In Bernd Enders (ed), *Neue Musiktechnologie II,* (pp. 216-229). Schott: Mainz.

10. Malamusi, M. A. (1999). From the Lake Malawi to the Zambezi. Aspects of Music and Oral Literature in South-East Africa in the 1990s. CD pamap 602 with booklet (27 pp.) Popular African Music/ African Music Archive: Frankfurt.

11. Schroeder, M. R. et al. (1979). Optimizing digital speech coders by exploiting masking properties of the human ear. *Journal of the Acoustical Society of America, 66(6),* (pp. 1647-1652).

12. Schumann, E. (1929) *Physik der Klangfarben.* Habilitationsschrift (ms.). Berlin.

ISBN 3-931852-67-9
ISSN 1617-6847