

GIORGIO ADAMO, Rom

Single Tone Colour and the ‘Sound’ of Music Items

Some remarks on the qualitative aspects in music perception

In music analysis we perceive a significant gap between the possibility of analyzing and describing single tones or sound phenomena and the difficulty of dealing with qualitative aspects of complex music sequences and items. The present paper will examine this matter with reference to the literature and through audiovisual examples, discussing some theoretical and methodological aspects such as the need of upgrading our conceptual and terminological tools.

Introduction

As an example for the definition of timbre let us consider the following typical remarks:

The American National Standards Institute (1973) defines [...] timbre as “... that attribute of auditory sensation in terms of which a listener can judge that two sounds, similarly presented and having the same loudness and pitch, are different”. Timbre therefore is defined in a purely negative manner as “everything that is *not* loudness, pitch, or spatial perception”. [...] The timbre of a sound, being itself a multidimensional attribute, depends on several physical variables. There is, in the first place, the frequency content and the spectral profile of the sound. [...] The temporal envelope of an instrumental sound, including attack, decay and modulation of the steady-state portion, influences the perceived timbre to such an extent that changes on any of them can make the sound of an instrument unrecognizable (Houtsma 1997).

Thereby, we can see how the concept timbre, or tone colour, has been used to identify a qualitative/multidimensional feature of single sounds, most frequently single tones. The research on timbre, consequently, was devoted mainly to the perception of single tones. But music, as we know, has to do with sound patterns within the time dimension. Basic musical concepts such as melody, harmony and

rhythm refer to sequences of sounds. In order to reduce the complex relation between the physical, the psycho-acoustical and the musical domain to the basic facts I will use the following chart:

Physics	Psychoacoustics	Music
Frequency	Pitch	Melody / Harmony
SPL	Loudness	Dynamics
Duration + SPL+	(Accent ?)	Rhythm

If we regard the timbre, we will get the following result according to this scheme:

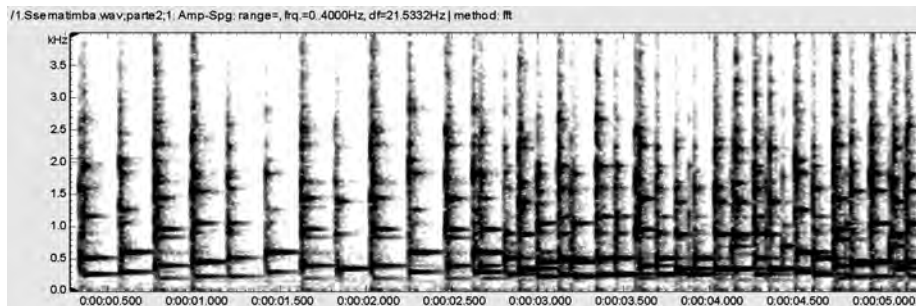
Physics	Psychoacoustics	Music
Spectrum + envelope	Timbre	???

The first chart shows how melody and harmony can be considered as a complex system (organization) of pitches, dynamics as a complex system (organization) of different perceptual intensities, rhythm as the result of the complex interaction of actually all sound parameters, starting with duration and SPL. The second chart shows that we miss a musical concept which refers to the complex system (organization) of timbre; we miss, in fact, concepts which combine qualitative aspects of both, vertical superimpositions of different sounds – for example the orchestral sound – as well as the horizontal structure (organization) of the time dimension, that is the timbre sequence.

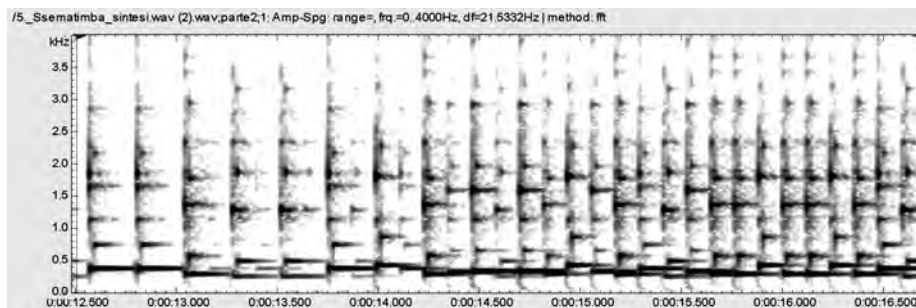
Timbre in the time dimension: stream segregation

The research on the role of timbre in stream segregation is a significant example for dealing with timbre in the time dimension. Albert Bregman (1990) and many others have demonstrated the role of timbre in Auditory Scene Analysis. Timbre similarity can be a decisive factor for grouping mechanisms which support the perception of separate sound sequences within a complex polyphonic music item. The general concept is clear and widely accepted even if how to define and estimate similarity/dissimilarity is still a research topic. In order to give an idea of the problems which can emerge from an experimental approach to such a complex parameter I will consider the sound examples Bregman offers in the CD *Demonstrations of Auditory Scene Analysis* (Bregman & Ahad 1995). For a sound demonstration of timbre segregation against pitch segregation Bregman has used *amadinda* music with reference to the study by Ulrich Wegner (1993). *Amadinda* music and the phenomenon of inherent – or subjective – melodic patterns, arising in the course of perception as a consequence of the interlocking technique used by two xylophone players, was clearly explained by Gerhard Kubik (for instance Kubik 1983, which includes both, analysis and transcription of *Ssematimba ne Kikwabanga*). This phenomenon can actually be considered as a perfect demonstration of pitch segregation. But the artificial demonstration of timbre segregation, as recorded on the mentioned CD doesn't seem rigorous enough: using synthesized sounds that imitate the *amadinda* xylophone and varying their spectral components Bregman paradoxically arrives at a good point to demonstrate this, even by a disputable procedure. The idea, as explained by both, Wegner and Bregman, was to interlock two sequences with the same pitch range and a different t i m b r e but I wonder to what extent we can consider the two synthesized interlocked sequences as having the same pitch range. In such a case it seems very difficult to separate timbre perception from pitch perception, especially, since we have a very strong spectral separation while we have to do with inharmonic sounds¹ (figure 1c).

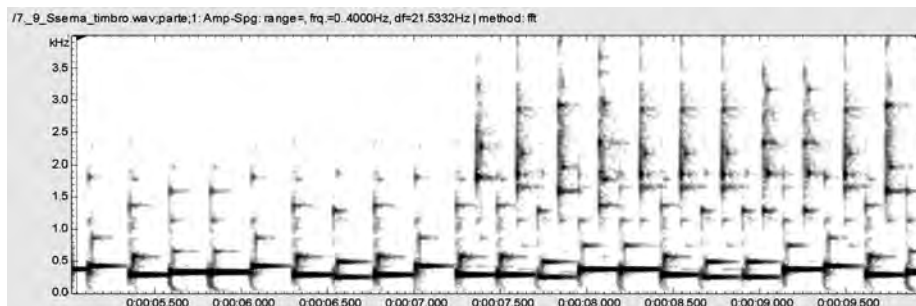
¹ For the problem of pitch perception in case of idiophones see Deutsch & Födermayr (1986).



a) *Ssematimba ne Kikwabanga*, original recording



b) Synthesized version



c) Synthesized version with “timbre” difference

Figure 1: Comparison of digital sonograms of about 5 seconds around the entry of the interlocked second part. Analysis made by the author on sound examples 7, 8 and 9 in Bregman and Ahad (1995).

Qualitative features as integration of sound parameters

We are getting here to the core of my remarks. In sound analysis and psycho-acoustical research we tend to use necessarily an analytical approach, that is, we try to isolate single aspects which can be significant for our sound and music perception. But our sound and music perception, at high hierarchical levels, is synthetic. We can try to plan laboratory experiments in order to isolate the effect of variations of a single parameter, but the actual music perception in the real world, within socio-cultural and historical contexts, is most of the time much more complex, in terms of the number of variables and their processing by the ear-brain complex, than our actual knowledge of the involved mechanisms. Coming back to the synthesized *amadinda*, it is clear that the stream segregation which occurs in the last of the three examples in figure 1 is determined by the change in the quality of the sound, but I wonder to what extent we can consider this change as a change in the timbre. From this point of view the concept timbre seems too complex, as it includes multiple dimensions (spectral distribution of the energy, inharmonicity, envelop) and at the same time too simple in comparison with the synthetic perception of all the parameters, including pitch and loudness.

Now, I would like to show a video example of the recordings I made in Malawi in summer 2008 during a video report about music and dance traditions in central and southern regions of Malawi that I carried out, accompanied – or better guided – by Dr. Moya A. Malamusi. After I had filmed a sequence of performances on *ulimba* xylophone accompanied by women's dance, Moya played for my documentation the "scale" of the instrument. I found it very interesting how the audience reacted to some "inconsistencies" in the sequence of sounds (*track 27* on DVD).

As we see in the video the xylophone is equipped with some kind of special slat which produces a sound completely different from the other ones, in practice it produces a percussive noise with no tuning. Thereby, we can observe the following phenomena:

- the most evident reaction is the laughing of the audience, especially the children, when the special slat is struck; we can see in figure 2 the difference in the digital sonogram between that slat – which is producing only noise with no recognizable partials and a very fast decay – and the sound of the adjacent keys (each slat is struck three times);

- the sequence of sounds is artificial and probably completely unusual within this culture; nevertheless, it seems to create a kind of expectancy, a pattern, whose denial provokes the reaction of laughing;
- in a normal performance, which includes three xylophone players, a drum and rattles, the sound of the specific slat is merged into complex sequences of different sounds and it is probably not even noticed by the audience. In fact, it produces a rhythmic pattern which is in the perception separated from the melodic lines of the xylophone (*track 28* on DVD, figure 3);
- the musicians do not show any particular reaction listening to the scale sequence; they already know the sound of that slat and there is nothing sounding strange to them;
- one could also notice that the children started to laugh lightly before that special slat was struck, in particular, after the last but three before the special one; I am tempted to explain this light laughing as a reaction to the differences occurring in the sound quality while descending in the scale: as we see in the sonogram (figure 4), we have a group of slats with good resonators, the sound of which is characterized by a wide range of harmonics, followed by a group of at least three slats with less partials, in non-harmonic relation to the first partial, and with a steeper decay in the rms curve.

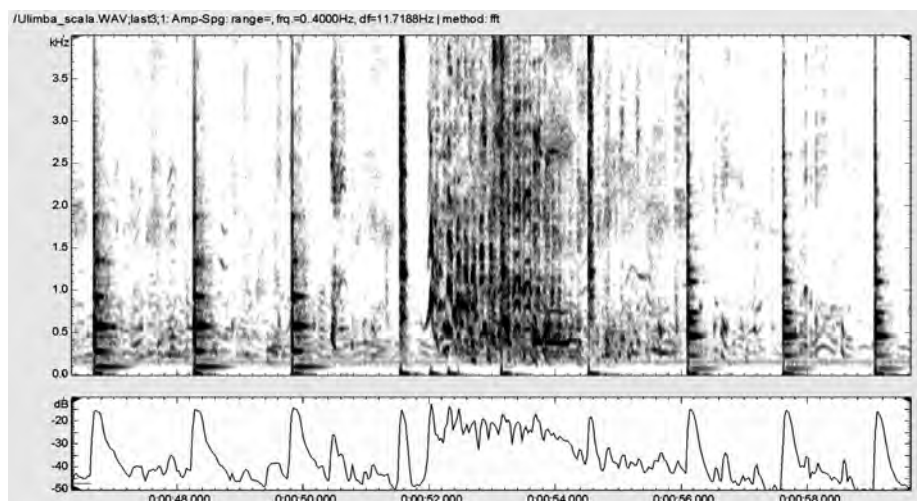


Figure 2: Last three slats in the descending scale, each one struck three times (see video example, *track 27* on DVD).



Figure 3: A stroke on the special slat during a real performance (frame of video example *track 28* on DVD).

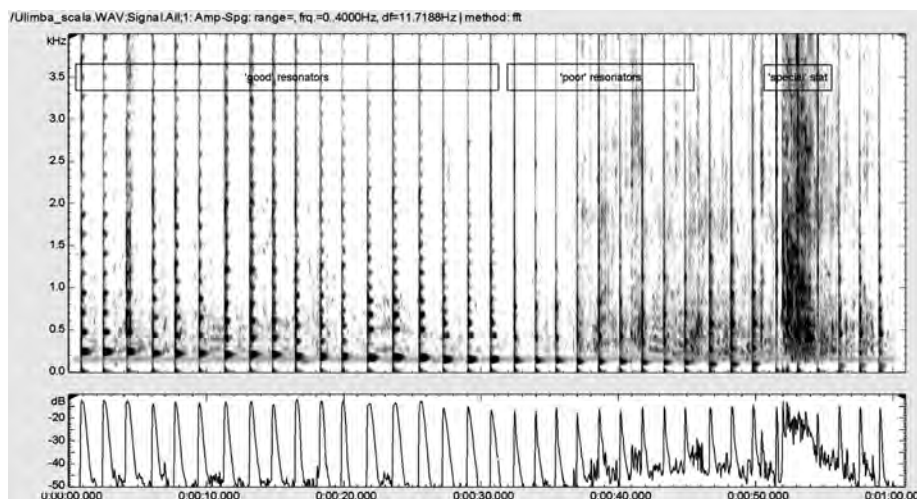


Figure 4: Descending scale (from video example, *track 27* on DVD), each slat is struck three times.

Comments:

- the context – in this case, the sequence – proves once again crucial for the musical perception; in a real performance (video example, *track 28* on DVD) the slat with no tuning gives rise to a consistent separated rhythmic pattern, while the two groups of tuned slats identified in the sonogram seem to lose their difference.
- it seems hard to explain exactly – like in the case of the synthesized *amadinda* – which parameter is primarily responsible for the perception of difference when listening to the scale. Actually we have a difference in the whole image of the sound. The quality as a whole is different. To what extent is the use of the concept of timbre satisfactory here?

The need of looking for more complex and at the same time flexible concepts referring to qualitative aspects of musical sound becomes stronger when we proceed from single sounds or small sound sequences to actual music performances. A nice demonstration of how synthetic our perception, memorization and recognition of musical sound actually is, was offered by Mantle Hood in a lecture given in Rome in 1992 (see Hood 1993). Dealing with *The untalkables of music* and referring to Hawaiian music he stated:

Identification of the sound of an individual singer or steel guitarist is made almost at once by the afficiado. This is really not too surprising. We remember the voice of an individual speaker on the telephone, even though we may not have heard the voice for some years. This audible "blueprint" can easily show the difference in partial spectrum found among singers and instrumentalists. Such a blueprint can even identify a whole unique tradition of music. (1993: 139)

At this point he played 25 "two-second wonders", i.e. a selection of very short sound examples from the most different music traditions², introducing them by the following remarks:

Some years ago I asked the 30-some members of the main seminar at UCLA [...] to select from their respective areas of expertise the briefest musical example that would identify the music. [...] The result was those 25 examples [...] each of them two seconds or less in duration. It was clear to everyone that the usual terms of musical analysis were not needed in the identification of the examples, e.g., meter, rhythm, harmonic progression, modal patterns, intervallic structure, scale, and so forth. (1993: 139-140)

Unfortunately, I don't have the original sound examples played by Hood but I will try to present something similar. To get to the point it is not necessary to compare the most different music ensembles. I can use two examples of different south-Italian vocal styles. These examples, in my opinion, clearly testify, how all the parameters contribute to building up the full image of the sound. Some years ago I published a thorough analysis of these two music items which can be considered as emblematic examples of two different musical cultures documented in South Italy in the '50: the rural musical culture of peasants and herders and the one of urbanized artisans living in small villages (Adamo 2002). On that occasion I used quite a number of pages and sonograms in order to demonstrate how the consistency of different parameters such as pitch range, melodic

² List of examples as provided by Hood (1993: 139): 1. atumpan; 2. Lobi xylophone; 3. Ewe sohu ensemble; 4. Sundanese gamelan; 5. Sundanese female singer; 6. Sundanese kecap; 7. Balinese gamelan gambuh; 8. kora; 9. Scottish bagpipe; 10. Ga (square-drum) ensemble; 11. Ibo master drum ("pot" drum); 12. disco (Nile Rodgers and Bernard Edwards); 13. pi'pa, "Snow in the Sunny Spring" (Lui); 14. Balinese gender wayang; 15. flute, tambura, mridangam, southern India; 16. Tibet funeral horns, procession, chant; 17. Korean kayagam; 18. Japanese gagaku; 19. vina (Nagaswara Rao); 20. Irish tin whistle; 21. Yugoslav balalaika orchestra; 22. organ, "Art of the Fugue"; 23. tenor sax (Johnny Beecher); 24. Bluegrass, "Faded Love"; 25. pansori singer and changko (Kwon-jin Chong).

contour, loudness, formant frequencies, waveform envelop, vibrato, etc. contributed to create the musical identity of the two items. It is interesting to find out how listening to a very short excerpt of three seconds and even one second reveals the essence of its identity (sound example, *track 31* on the DVD).

In figure 5 we can see the digital sonograms of the three-second portions of each item. The graphs quite clearly show the particular combination of various parameters which characterize each performance in a very contrastive way: fundamental frequency range, distribution of the energy within the spectrum, handling of intonation, presence/absence of glissando and vibrato, dynamics, duration of single tones, sharpness of attack and decay. It is actually the integration of all these parameters, at an acoustical, psycho-acoustical and musical level, that identifies the two items even in a very short excerpt of only one second.

The problem of extracting qualitative features from music items is not new. In the case of vocal music the research has taken advantage of phonetic studies, methodology and terminology regarding the voice quality (e.g.: Laver 1980, paradigmatic in its title). The lifelong work of Johan Sundberg on the singing voice, and the many contributions by Franz Födermayr in the analysis of vocal styles can be considered as significant examples of the close link between phonetics and musicology in this field. In general, it seems easier, when the voice is involved, to speak in terms of qualitative aspects (falsetto, nasality, singing formants, etc.) that are recognizable in the sound analysis and at the same time near to the actual sound and musical perception.

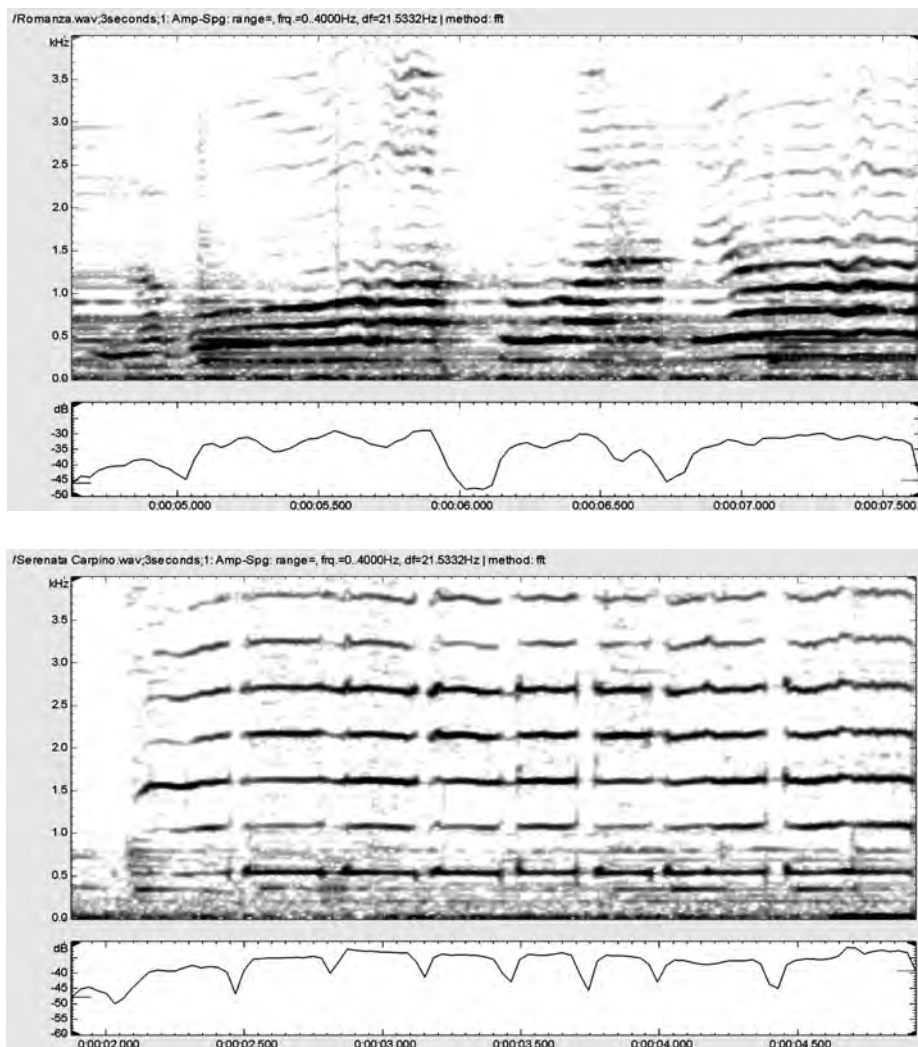


Figure 5: Three-second excerpts of two songs recorded 1954 by Alan Lomax and Diego Carpitella in two villages of Apulia (South Italy): *Romanza*, male voice and guitar, village of Sannicandro (above), example of the style used by urbanized artisans; *Serenata*, male voice and *chitarra battente*, village of Carpino (below), example of rural style. Original sound examples from Lomax and Carpitella (eds.) 1957.

The concept of Klangideal, introduced by Franz Födermayr (1971: 101, 146-154) represents an interesting attempt to deal with the highly complex synthesis of quantitative and qualitative aspects which belong to the actual musical perception within a given culture. I have found this concept very fascinating and I have used it in some ethnomusicological studies of Italian music traditions (e.g. Adamo 1991: 225; 2005: 158).

The Long-Time-Average-Spectrum, introduced in music analysis by Jansson and Sundberg (1972) and used for example by Födermayr, Deutsch and Jena (1992) in a study of the sound of orchestras, is a method of sound analysis and visualization which has been in my opinion a step forward in the direction of extracting qualitative features from complex music items. I have used LTAS successfully in some cases, for example, in order to show clear similarities between vocal and instrumental sounds (Adamo 1992), or in a comparison between spoken and sung text (Adamo 1994). I wonder why such a method has not gained a wider application within the study of qualitative aspects of music. It sometimes seems to me that the conceptual tools used by musicologists and the ability of the latter to identify targets for analysis are less effective than the methods and techniques at their disposal.

Need for widening musicological concepts and terminology

Assuming that the research on timbre is of greatest importance for both, psychoacoustics and musicology, is it possible to try to go further in the investigation of qualitative aspects in music? Let us examine an example of possible future directions here.

The way different types of instruments are combined in music traditions can be seen as a way of combining different *timbre patterns* (specific to each type of sound producing instrument and/or technique) in order to produce specific complex *timbre configurations*, for instance voice + plucked string, fife + drum, etc. During my field research in summer 2008 in Malawi I recorded various so-called banjo groups: within these ensembles one can find an interesting variety of home-made instruments which are sometimes quite dissimilar but can be referred to the general configuration of voice(s) + guitar(s) + bass + drum. In this context, the basic drum set of rock and jazz music, in particular, is reinterpreted in order to maintain some contrastive sound features, imitating somehow the sound quality of the hi-hat, toms, bass drum, crush and

ride cymbals etc. This kind of imitation is clearly audible in the music of the *Muonekela Band*, a small group of street musicians recorded in Blantyre (figure 6 and video example, *track 29* on DVD). It is interesting to notice similarities and differences when we examined another group, for example the *Madetsa Band* (video example *track 30* on DVD). In this case, one singer plays the bass drum (actually the sound box of the bass string instrument) by hand sitting behind the banjo players and beside the bass player while the ingenious percussion set in the foreground (figure 7), with its small bells and some kind of metal drum, seems to show influence of the bell-drums combination widely spread in Africa.

What I'm calling here *p a t t e r n s* and *c o n f i g u r a t i o n s*, in absence of a more precise terminology, is something probably relevant to both, the intra- and trans-cultural musical perception but quite far away from a satisfactory musicological description. In the field of popular music the term *s o u n d* has sometimes been used, maybe in a naive way, as reference to the general acoustic image of a group. To say that the sound of the Beatles is much different from that of the Rolling Stones, or that Jimi Hendrix has transformed radically the sound of the electric guitar can actually make sense. Similar terminology has been used in the world of the so-called soundscape composition. To what extent could we use or re-define this kind of categorization – far apart from a “scientific” terminology but very near to the experience of the auditory – in order to develop new conceptual tools and focus new fields of research?

After all, can we try to move forward by filling the gap between the colour of the single tone and the qualitative aspects of music items?



Figure 6: *Muonekela band*, recorded in Blantyre, July 23rd, 2008. Frame of video example, *track 29* on DVD.



Figure 7: *Madetsa Band*. Recorded in Madetsa Village, T.A. Mazengela, Lilongwe, July 24th, 2008. Photo by Moya A. Malamusi. (Video example, *track 30* on DVD)

References

- Adamo, Giorgio
 1991 "Musica come evento sonoro: analisi acustica di canti a zampogna della Basilicata." In *L'analisi musicale*, ed. by R. Dalmonte & M. Baroni, 221-240. Milano: Unicopli.
- 1994 "Metrica cantata, metrica recitata." In *Il verso cantato. Atti del Seminario di Studi (aprile-giugno 1988)*, 55-68. Roma: Università di Roma La Sapienza.
- 2002 "La voix et l'identité culturelle: un regard sur les traditions du Sud de l'Italie." In *La vocalité dans les pays d'Europe méridionale et dans le bassin méditerranéen*, ed. by L. Charles-Dominique & J. Cler, 129-142. Saint-Jouin-de-Milly: Édition Modal.
- 2005 "Aspekte der Zweistimmigkeit in der Volksmusik Italiens." In *Mehrstimmigkeit und Heterophonie. Bericht zur Tagung in Wien 11. bis 12. Dezember 1999*, ed. by G. Gruber, A. Schmidhofer & M. Weber, 153-167. Frankfurt am Main: Peter Lang.
- Bregman, Albert S.
 1990 *Auditory Scene Analysis. The Perceptual Organization of Sound*. Cambridge (Mass.): The MIT Press.
- Födermayr, Franz
 1971 *Zur gesanglichen Stimmgebung in der außereuropäischen Musik*. Wien-Föhrenau: Stiglmayr.
- Födermayr, Franz, Werner A. Deutsch & Stefan Jena
 1992 "Digitale Langzeitanalysen als neue Methode der Untersuchung von Orchesterklängen." In *Klang und Komponist. Ein Symposium der Wiener Philharmoniker*, 311-326. Tutzing: Hans Schneider.
- Hood, Mantle
 1993 "The untalkables of music." *Annuario degli Archivi di Etnomusicologia dell'Accademia Nazionale di Santa Cecilia*, 1: 137-142.
- Houtsma, Adrianus J.M.
 1997 "Pitch and timbre: Definition, meaning and use." *Journal of New Music Research* 26, 2: 104-115.
- Jansson, Erik V. & Johan Sundberg
 1972 "Long-time-average-spectra applied to analysis of music", *STL-QPSR*, 13, 4: 40-45.
- Kubik, Gerhard
 1983 "Die Amadinda-Musik von Buganda." In *Musik in Africa*, ed. by A. Simon, 139-165. Berlin, Staatliche Museen Preussischer Kulturbesitz – Museum für Völkerkunde.

Laver, John

1980 *The phonetic description of voice quality*. Cambridge: Cambridge University Press.

Wegner, Ulrich

1993 "Cognitive Aspects of *amadinda* Xylophone Music from Buganda: Inherent Patterns Reconsidered." *Ethnomusicology* 37, 2: 201-241.

Sound Records

Bregman, Albert S. & Pierre A. Ahad

1995 *Demonstrations of Auditory Scene Analysis: The Perceptual Organization of Sound*, CD, MED 9535. McGill University: Department of Psychology, Auditory Perception Laboratory.

Lomax, Alan & Diego Carpitella (eds.)

1957 *Southern Italy and the Islands – World Library of Folk and Primitive Music vol. XVI*. 1 LP 33rpm 30cm, Columbia KL/5174.

List of examples on the DVD

Video:

Track 27: Scale played by Moya A. Malamusi on ulimba xylophone (Samalamoyo Band). Khungubwe Village, T.A. Ngowe, Chikwawa District, 17.07.08.

Track 28: Kahombe (mwana). Samalamoyo Band: xylophone (ulimba), drum, rattle and singing; ethnic group: Sena. Place and date as example 1.

Track 29: Kuli milandu. Muonekela Band: Tingo Makwinja, guitar, Vincent Makwinja, bass, Mcnight Waliya, drums; ethnic group: Lomwe. From Lukula Village, T.A. Chimaliro, Thyolo District, but rec. in Blantyre, 23.07.08.

Track 30: Usandi namize. Madetsa Band: four banjos, bass, drums; ethnic group: Chewa. Madetsa Village, T.A. Mazengela, Lilongwe District, 24.07.08.

Sound:

Track 31: A/B comparison of 3 seconds (two times) and 1 second (two times) excerpts of two recordings collected in South Italy by Lomax and Carpitella: Romanza, voice and guitar, Sannicandro (Foggia), 22.08.1954; Serenata, voice and chitarra battente, Carpino (Foggia), 23.08.1954. (Lomax and Carpitella 1957)