Udo Will

An ethnomusicologist's dilemma: Technology and the analysis of music

The present century has seen an extraordinary development in laboratory instrumentation in many a discipline, due to the advances in technology in general and in electronics in particular. In several disciplines these developments have led to or at least contributed significantly to the emergence of new paradigms that have changed our views of the world. Astronomy and cosmology, particle physics and molecular biology as well as fields like mathematics (chaos theory) and theoretical physics have experienced drastic changes and developments under the influence of new laboratory technology, and it seems justified to say that without the backing and often the stimulating initial discoveries due to new instrumentation, our interpretation of the world would be essentially different to day.

Once they have been invented many of these lab instruments have also been put to use in disciplines other than those they were originally created for. Musicology has always had its share in adapting and transforming the emerging instrumentation for its own purposes, especially instruments for recording and analysing acoustical events, from optical mirror oscillographs to analog recorders and the latest computerized digital sampling and analysis units. Remarkably, though, and unlike what has happened in other disciplines, there seems to have been no comparable accompanying development in the way we analyse music. It can be argued that, apart from the approach of the remarkable group of scholars around the Seashores, Harold and Carl, that is, there have been no paradigmatic changes in analytical methodology complementing and advancing the possibilities opened up by technological developments (Seashore, 1938, 1947).

The main musicological fields in which modern (computer-) technology has gained some importance are documentation and analysis of notated music. Already in 1967 Bengtsson had put attention to the fact that the whole area of analysis of 'music as performed' has been, by and large, unaffected by these technological developments – and nothing much has changed since then. It appears that a major reason for this situation is to be seen in a historical legacy of musicology, the fact that it arose from within the context of Western musical culture and developed notation and the analysis from notation as its standard and dominant methodological approach. So engrained was and still is this methodological orientation that different approaches of analysing music by technical means were always conceived of as forms of 'notation', as is obvious from terms like 'automatic music notation', 'objective music notation', etc. As Mantle Hood has put it, the application of this instrumentation was nearly exclusively determined by "a desire to make visible to the eye illusive refinements of sound readily perceived by the trained ear" (M. Hood, 1993, p.138), meaning that these instruments were not primarily conceived as analytical tools but as means to complement standard notation through visualisation of musical elements not notatable otherwise.

As sophisticated as for instance the different stages in the development of Seeger's melographs have been – always at the forefront of the current state of technology – he somehow stopped short of making available the full potential of his machines by focusing mainly on 'the development of the graph' i.e. graphic registration (Seeger,1958). Musical analyses performed with the aid of the melograph or similar instruments are all confined to a more or less qualitative description of the output (fundamental frequencies are estimated by comparison with a grid of reference lines, frequencies of harmonics are determined via the frequencies of fundamentals, etc.) and if quantitative data are supplied they are estimations with no indication of the degree of accuracy or the method by which they were obtained (e.g. Morton,1974 and other contributions in vol.II, no.1 of the 'Selected Reports in Ethnomusicology' which focuses on analyses performed with the Melograph Model C). This limitation, however, was not imposed by the available state of technology as other disciplines in fact had developed means for quantitative evaluations of data from similar instruments.

It is worth recalling that notation has not arisen as an analytical tool but as a memory and organisational aid within and for the music of a specific culture. If we (mis-)use notation to analyse music, we have a situation C. Seeger has pointedly characterised in his article on prescriptive and descriptive music writing: 'In employing this mainly prescriptive notation as a descriptive sound-writing of any music other than the Occidental fine and popular arts, we do two things, both thoroughly unscientific. First we single out what appears to us to be structures in the other music that resemble structures familiar to us in the notation of the Occidental art and write these down, ignoring everything else for which we have no symbols. Second, we expect the resulting notation to be read by people who do not carry the tradition of the other music. .... To such a riot of subjectivity it is presumptuous indeed to ascribe the designation 'scientific' ' (Seeger, 1958, 186-187). The use of conventional notation for analytical purposes has to be considered as one type of classification of perceived sound events according to a specific reference system. First, as we are dealing with classification of *perceived sounds*, this method obviously has a strong, unavoidable cultural bias caused by the perceptive-cognitive system of the transcriber. Secondly, classification according to a specific reference system (standard notation of Western music) introduces a further bias: only those sound events that are referenced by the accepted symbol system can be notated. With notation the actual musical behaviour has gone through a complex physiological and psychological filter and has correspondingly been transformed. Nevertheless, this may be fine and acceptable in certain cases as long as the filter of those who produce the behaviour and those who analyse it are of the same kind (the musician as musicologist; or both are different individuals but come from the same culture) and as long as no wrong claims are made about what actually has been analysed. This is however not the way to go if we want to analyse music from a different culture because analysis from notation is necessarily always analysis of 'music as perceived (by the transcriber)', not of 'music as performed'.

Compared with notation, automated methods obviously offer a more direct and less biased approach for the analysis of musical behaviour. For one, we are no longer dealing with perceived sound events. Electronic registration can assist in the analysis of what happens on the signal side and provide information about events that are inaccessible or go unnoticed in the standard notation approach due to the filter characteristics of the cognitive apparatus of the transcriber. I think there is not much to discuss about the status of automated registrations: they can give a description of certain parts of a particular series of physical sounds, i.e. of components of a particular performer's behaviour at a certain moment. This is well and good; but difficulties

begin when we start to read and interpret these registrations. It is often agreed that when confronted with the output of automated analysis, on does not see the wood for the trees (e.g. Jairazbhoy, 1977). But here I would side with Mantle Hood who said that '.....the problem is never one of *too much* information, but rather what does the information mean? How do we evaluate it?' (M. Hood, 1993, p.138).

I want to submit that this evaluation has to proceed in two steps to render a musically meaningful analysis: we have to take into account the physiological charac-teristics of the human auditory system and we have to consider it's psychological-cognitive features. If electronic registration results in a description of certain parts of a physical sound series then it is obvious that not every analysable aspect of this registration is musically relevant. It may be a fact that two drum beat pairs differ in length by 12 msec or that two tones of 100 and 500 Hz have amplitudes of 50 and 40 dB respectively, but due to the physiology of the auditory system, these differences are not perceived as such and have no musical significance. Therefore, the first necessary step must be a data reduction on the basis of auditory physiology. Today psychophysics has accumulated a body of knowledge large enough to have this step resulting in a *perceptually* meaningful reclassification of the original sound data. A subsequent second step has to aim at an analysis of the psychological-cognitive features in order to arrive at a *musically* (and culturally) significant description of the recorded musical behaviour. Fortunately, we do not need any prior knowledge about the underlying cognitive system in order to perform this task, because production of musical behaviour is a feedback process : the output (musical behaviour) is under constant control of the performer who can immediately correct the output if necessary. Because of this continuous readjustment, an adequate analysis of the output allows to determine the underlying control criteria. To give a very basic example, it is impossible for a singer to produce a tone of strictly constant frequency. He has constantly to re-adjust the frequency and controls the output via two feedback loops, 'internally'(i.e. inside the performers body) via the re-afferents from the larynx musculature and externally via the acoustic feedback to the ear, resulting in an oscillating frequency output. Although the reason for this oscillation is the physiological impossibility to produce a constant vibration of the vocal cords, the extent to which the frequency is allowed to vary has a cognitive aspect: it depends on what is considered an acceptable variation in tone production (pitch class) in the respective culture – and this variation can be determined precisely by analysis of an adequate sample of registrations.

In the following I am going to exemplify this approach with two examples, the first of which concerns frequency and pitch measurements. Analyses start with determining the period duration and/or the frequency of every cycle of the registered sound wave item (a song verse, for example). These measurements, which, if plotted as a time-frequency graph, represent a graphical notation of the actual frequency production of the item concerned, contain information that is perceptually not very meaningful. Psychophysics tells us that there is a time constant working with pitch perception, i.e. a minimum time and/or a minimum number of soundwave cycles is required for the perception of pitch. This integration time of about 20-25 msec is, however, impractical to use. For vocal music with a frequency range of about 100-400 Hz this difficulty can be circumvented by averaging the original measurements over 3 to 5 cycles, resulting in a somewhat 'smoothed' data set which then can be analysed. There are several ways in which the frequency content can be analysed and the actual methods have to be chosen according to the type of melody: A melody with clearly separated tone or pitch levels (like, for example, Thai vocal music) invites for a different type of analysis than a melody with more gliding pitch levels. An example of the latter type is vocal music

from Central Australia, for which frequency measurements from one example are shown in the upper diagram of fig.1.
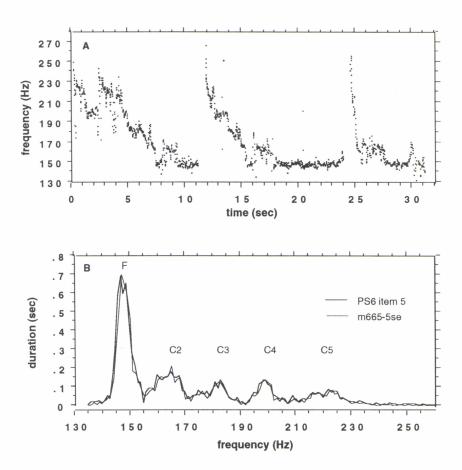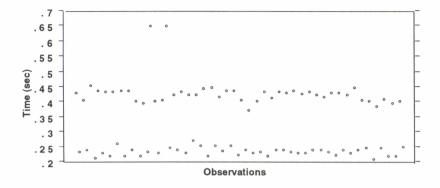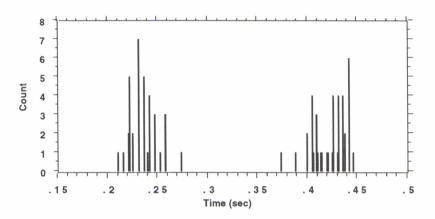

Figure 1.

For this type of melodies is has been proven useful to analyse the frequency content in terms of a frequency-duration graphs: All measurements are grouped in classes (e.g. 1/4 Hz or 1/2 Hz classes) and the number of occurrences for each class is determined; the duration of each (center-)frequency can than immediately be calculated. The basic idea behind this is, as has been shown by psychophysical experiments, that the longer a frequency is produced during the performance of an item the more relevant it is for pitch perception. The frequency duration graphs (like lower graph in fig.1) show a series of crests (five in the given example) of which most have a clear center or main frequency peak. Subsequently, the combined analysis of the time-frequency and frequency-duration data allows to determine dominant frequencies, stability of intonation, intervals of the item, etc. (see e.g. Will and Ellis 1996) The significance of this information increases if we analyse more than one item. The lower graph of fig.1 shows the frequency-duration plots of two items (two verses with the same text) and it is immediately obvious that not only the frequencies are nearly the same (deviations do not exceed ± 1Hz) but also the durations are nearly perfectly reproduced. These results

demonstrate not only a very precise and stable text-melody interlocking but also an extraordinary perfection of frequency production and, consequently, an extreme sense of pitch of Australian singers. The stability of frequency production can be understood as an expression of the narrowness of the pitch categories in the music concerned. In other cultures, for instance Western Music, pitch categories are much broader and frequency production shows considerably larger variability. Performing such an analysis for many more items with different pitch levels, we are also able to reveal underlying interval concepts. For example, the analysis of more than 50 song items from Central Australia has given evidence that Aboriginal singers use an interval system based on frequency differences, whereas almost all other musical cultures use frequency ratio intervals (Will, submitted). Furthermore it could be shown that Aborigines do not apply the principle of octave identity to the construction of their melodies although they do make use of an isolated octave ratio interval in some of their songs.

My second example concerns the analysis of a specific rhythmic figure – a double-beat pair – figuring prominently in Central Australian clap stick accompaniment. If we measure the timing of the clap-stick beats by the points of their maximum amplitude than a plot of the beat intervals results in figures like the upper diagram of fig. 2: there are two types of intervals a shorter one of about 0.2-0.28 sec and a longer one of about 0.37-0.45 sec (the two 'outliers' of about 0.65 sec are two single beats being performed instead of the double-beat pairs). Re-grouping the measurements in 1 msec classes (the original recording had been sampled at a rate of 19.2 kHz) and plotting these measurements in histogram form the grouping into the two types and the distribution within each group are obvious (middle diagram of fig. 2).
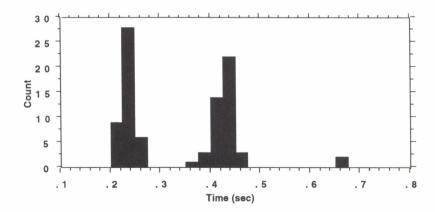
Figure 2.

However this histogram still contains a lot of information that is physiologically and psychologically not relevant and that can be further reduced by the following considerations:

1) The time resolution of the (peripheral) auditory system is about 3 msec. This means acoustic events must have at least a time difference of 3ms in order to be perceived as different events.

2) Acoustic events separated by a time difference larger than 3 msec and smaller than 20-30 msec do have a space direction but not a time direction, meaning that although we perceive them as coming from certain directions of the surrounding space, we are not able to tell which is the first and which the second. This resolution limit is probably caused by more central processes because it is basically the same for different modalities (acoustic, visual, tactile).

Taking these two different time resolutions into consideration it seems justified to re-group our measurements in 25-30 msec classes – differences smaller than this will be, if at all, of little perceptual and cognitive relevance. (One could even argue that in our example the 'time resolution' is likely to be coarser as we are dealing with rhythmic or interval sequences, i.e. comparison of successive and not 'simultaneous' acoustic events.) With 25 msec classes we obtain the lower diagram of fig.2. The shorter interval shows a nearly normal distribution with a center-class mean of 0.2375 sec and the longer interval has a mean of 0.4375 sec. The distribution of the latter interval seems slightly skewed to the left, but analysis of the complete song is needed to tell whether this is of any significance (it is possible that the skewedness is caused by some intervening factors like alignment of accompaniment with melody or text). For the analysed song section the average ratio of the small interval (double-beat) to the whole figure is 0.358 and the figure is performed at a speed of 90.56 figures/minute.

The above rhythmic figure has been notated and analysed by various researchers in different ways and there is a continuing debate over which notation is the 'right' one. Two frequent notations are 1) dotted eighth note and a sixteenth note tied to a quarter note and 2) eighth note and a quarter note. The ratio of the shorter interval to the whole figure is 0.375 in the first notation and 0.333 in the second. For the analysed song section with a ratio of 0.358, however, both notations are wrong and do not describe the actual rhythm performed which is extremely difficult to notate precisely in western notation. Although it can be argued that both notations can equally serve as approximate notations (at the given tempo, both miss the second beat by less than 20 msec), they do not help understanding this rhythmic figure. Obviously, the above analysed rhythm does not match simple triple or quadruple subdivisions of the whole figure and a preliminary analysis of a large sample of songs from Central Australia indicates that that this also hold for other songs. Rhythmic structures in this music do not seem to follow structures reflected in Western notation. It is severely misleading to analyse and understand this figure from an oral culture in terms of rhythmical concepts evolved in and for a literal culture. The Central Australian double-beat figure does not arise from simple triple or quadruple subdivisions of the whole figure and its origin must be sought somewhere else. An interesting explanation of this rhythm has been attempted by R. Moyle who tried to link this figure to the human heart beat sequence (Moyle, 1995). Whether and to what extent this explanation holds can, however, only be shown by analysis of a larger sample of this accompaniment figure from Central Australia.

With these two examples I hope to have shown that analysis of electronic registrations of musical behaviour is in no way limited to a mere physical description of sound events. It even enables us to gain knowledge about perceptual and cognitive aspects of a performance (and the performer) which are often thought to be accessible only via the empathic mediation of a transcriber. Very likely the methods I have presented here

are not applicable in all cases, and there are several types of music that require completely different approaches. Unfortunately, as a consequence of the quasi ineradicable fixation on notation and 'graphical descriptions' musicology does not yet dispose of an established body of methodological approaches for the analysis of automated registrations. The Seashore group has done admirable work in this area and has indicated a direction to take, although some of its methods need revision today.

References

Bengtsson, I. 1967. On Melody registration and 'Mona'. in Heckmann, H. (ed) *Elektronische Datenverarbeitung in der Musikwissenschaft.* G.Bosse, Regensburg, pp. 136-174.

Hood, M. 1993. The untalkables of music. In: EM *Annuario degli Archivi di Etnomusicologie dell'Accademia Nazionale di S.Cecilia*, I, 1993, pp.137-142.

Jairazbhoy, N. A. 1977. The 'Objective' and Subjective View in Music Transcription. *Ethnomusicology*, May 1977, pp. 263-273.

Moore, M. 1974. The Seeger Melograph Model C. *Selected Reports in Ethnomusicology* Vol.II, No.1, pp. 3-13.

Morton, D. 1974. Vocal Tones in Traditional Thai Music. *Selected Reports in Ethnomusicology* Vol.II, No.1, pp. 89-99.

Moyle, R. 1995. Singing from the Heart? in: Barwick et al. (eds): The Essence of Singing and the substance of song. Sydney, Oceania Monographs No.46, pp.53-58.

Seashore, C. 1938. Psychology of Music. MacGraw-Hill.

Seashore, C. 1947. In Search of Beauty in Music. A Scientific Approach to Musical Esthetics. The Ronald Press Comp., New York.

Seeger, Ch. 1958. Prescriptive and descriptive music writing. *Musical Quarterly* 44, 2.

Shonle, J. I., and Horan, K. E. 1980. The Pitch of vibrato tones. *Journal of the Acoustical Society of America* 67 (1): 246-252.

Will, U. and Ellis, C. J. 1996, A re-analyzed Australian Western Desert Song. *Ethnomusicology*, Vol.40/2, pp. 187-222.

Will, U. (submitted). Evidence for non-logarithmic frequency intervals in Australian Aboriginal Music. *Journal of the Acoustical Society of America* (submitted).