Music technology and audio processing: *rall.* or *accel.* into the new millennium?

JOHAN SUNDBERG

Voice Research Centre, KTH (Royal Institute of Technology), Department of Speech, Music and Hearing, SE-100 44 Stockholm, Sweden

Music acoustics research can provide support in terms of objective knowledge to further the rapidly developing areas of music technology and audio processing. This is illustrated by three examples taken from current projects at KTH. One concerns the improvement in quality of sound reproduction systems over the last century. A test, where expert listeners rated the year of recordings of different ages, demonstrated that significant advances were made between 1950 and 1970, while development was rather modest before and after this period. The second example investigates the secrets of timbral beauty. Acoustic analyses of recordings of an international opera tenor and a singer with an extremely unpleasant voice shed some light on the basic requirements of good vocal timbre. The unpleasant voice is found to suffer from pressed phonation, lack of a singer's formant, irregular vibrato and insufficient pitch accuracy. The third example elucidates tuning and phrasing differences between deadpan performances of MIDI files played on synthesizers and performances by musicians on real instruments. The examples suggest that future development in the areas of music technology and audio processing may gain considerably from a close interaction with music acoustics research.

1. INTRODUCTION

Many years ago, a man from the Stockholm Royal Opera called me and asked if I could help with the tuning of their bell synthesizer. This device, presumably more than twenty years old at the time, and enclosed in box of solid oak, turned out to be a piece of truly historic electronics. The tuning problem was due to the fact that the frequencies of the partials changed with their decaying amplitudes. A question then entered my mind: why should music culture content itself with using obsolete technology? The inevitable conclusion was that music acoustics must facilitate the use of contemporary technology in music. Hence, technical development is an important branch of music acoustics, particularly in Sweden where it relates to the Faculty of Electronics at KTH. Fortunately, the problem was solved when Sten Ternström appeared in the music acoustics class and expressed interest in constructing a new, modern bell synthesizer. He called it ISABEL, and it has since been used many times at the Royal Opera in Stockholm (Ternström 1981).

Another revelation struck me some years ago, when I witnessed the preparation for a pop concert. Lots of heavy boxes were carried from the truck to the stage. However, very few of them contained musical instruments, but rather electronic equipment. When the set-up was completed, the scene was crowded by loudspeakers, various twinkling boxes and a spaghetti of cables. This made me realise that the situation had now changed entirely. Previously, music relied on obsolete technology; today, modern technology has invaded music culture.

This invasion has brought some concerns. For example, the Internet offers downloadable MIDI files that can be played on the local software synthesizer of a personal computer. The resulting performance is usually deadpan, i.e. it replicates exactly what is written in the score. Real performances by living, educated performers contain a number of deviations from the score. These deviations have been shown to add significantly to that musical quality of the performance which makes listening worthwhile. One may ask, then, why the opportunity to listen to deadpan performances is offered. The reason would be that engineers have failed to realise the importance of expressive deviations in music performances. Had they asked musicians, they would probably have been informed that such deviations are mandatory, but the musicians' descriptions of the deviations may have been difficult to convert into a rigid computer program. The simplistic solution would be to go for deadpan performances.

Another concern has arisen in the area of synthesizer technology. The development in this area during the latest decades has been characterised by a quasi-periodic pattern of innovation, hope and despair. We have seen a series of technological breakthroughs, greeted by optimism, but after some years by a somewhat lethargic despair. For example, FM synthesis caused a revolution in synthesizer technology, and the Yamaha DX-7 almost killed off all competing devices. However, musicians soon felt that this synthesis method suffered from some serious limitations, and started to hope for something better. The situation today is, almost ironically, similar. The availability of massive computing power allows programmers to implement mathematical models of conventional instruments. At the same time, we do not know if all the details included in these descriptions really contribute relevant properties to the resulting sound quality.

Organised Sound **4**(3): 153–160 © 1999 Cambridge University Press. Printed in the United Kingdom.

Thus, while music culture is flooded by the latest technology, the need for a defence against oversimplification has grown. Today, it is generally realised that what has not been explicitly formulated may still be highly relevant. Music technology seems to need basic knowledge in many areas, e.g. definitions of how a performance of a piece of music should deviate from deadpan in order to sound musically interesting, or what the essential properties of music instruments must be, in order to be useful.

The discipline of music acoustics should be capable of providing this basic knowledge. Its key questions are the same as in other areas in the natural sciences: HOW? and WHY? The question HOW? requests descriptions of music sounds in acoustical terms. The question WHY? is more complex. Music sounds possess their specific properties for a variety of reasons. They must fit the human auditory system, so part of the explanation can be found in auditory perception. The construction and function of music instruments provides another type of explanation. Research in the area of music acoustics should therefore aim to describe what the important properties of a musical instrument are, and what is needed to make a musical performance worthwhile for the listener. Thus, it must address those questions that need to be answered for a successful implementation of modern technology in music culture.

Honouring Darwin, we may say that traditional instruments are the products of a long history of attempts at different methods and constructions for sound production. They have all been constantly matched against the rigid, though completely elusive demands of ‘musical usefulness’; only the musically fittest instruments have survived. What are the characteristics of this filter? To find the answer, it would be worthwhile to study musical instruments that have been thoroughly processed by this filter for a long time. The human singing voice would be a good candidate as it, unlike most other instruments, has been filtered through the demands of ‘musical usefulness’ for thousands of years.

In this article, I will describe three areas of music acoustics research that seem relevant to music technology and audio processing. They concern the development of sound reproduction, the characteristics of the singing voice, and the synthesis of music performance.

2. YEAR OF RECORDING

The quality of sound reproduction has increased dramatically in the last century, from Edison’s early attempts to today’s digital hi-fi. In an investigation of the perceived age of classical singers, we asked a panel of opera fans to rate not only the age of the singer, but also the age of the recording (Sundberg, Niska-Thörnvik and Söderström 1998). The guessed age of the singers showed a reasonably linear relation with their real age, both for male and female singers (figure 1(a)). It turned out that the frequency of the vibrato undulations of the fundamental frequency significantly contributed to guessed age; the slower the vibrato, the higher the guessed age.

The guessed recording year showed a relation to the real recording year that was nonlinear, and more similar to a cosine or a smoothed step function (figure 1(b)). This seems to imply that the perceived quality of sound recordings remained at a rather modest level from 1920 to about 1950, i.e. during the 78 rpm era, and then improved substantially between 1950 and 1970. During the last twenty-five years, however, not much seems to have happened, in spite of the advent of digital recording technology. The enhancements to sound quality apparently escaped the expert panel used in this experiment. The reason may have been that the excerpts were presented to the listeners on regular cassette recordings. Anyhow, even if the recent developments in the area of audio processing may be hard for many listeners to detect, it has certainly facilitated the work of the sound engineer.

3. SEARCHING FOR THE ESSENCE OF MUSICAL SOUND

The complexity of musical sound differs greatly between instruments, since the acoustic flexibility varies over a wide range. The pipe organ seems to represent one extreme, basically allowing nothing more than the binary choice of whether or not a pipe sounds at a certain moment. All other properties are in principle predetermined, and thus beyond the player’s control. The contemporary computer music studio represents the opposite extreme, permitting any type of sound property to change at any moment in time. The traditional instrument closest to allowing a similar freedom for the musician is the human singing voice. It imposes no restrictions regarding variation of loudness and tuning within the singer’s pitch range, and the restrictions on resonance frequencies are small. Yet, this great flexibility obviously does not necessarily lead to excellent instrument quality; it depends on the skill of the musician.

Research on the singing voice has focused primarily on excellent voices in the western operatic classical tradition. However, instructive examples are also provided by voices that are generally regarded as unpleasant. A well-known example is Miss Florence Foster Jenkins, documented on a CD (GD 61175, BMG Classics). She had an exceptional, though paradoxical career due to her astonishingly bad singing voice. The most salient aspects of her vocal shortcomings seem trivial, as they concern failures to reach target pitches.

The same CD contains recordings of quasi-tenor singer, Thomas Burns, who was also in possession of an extremely unpleasant voice. However, in this case, the mismatching of target pitches did not seem to be the

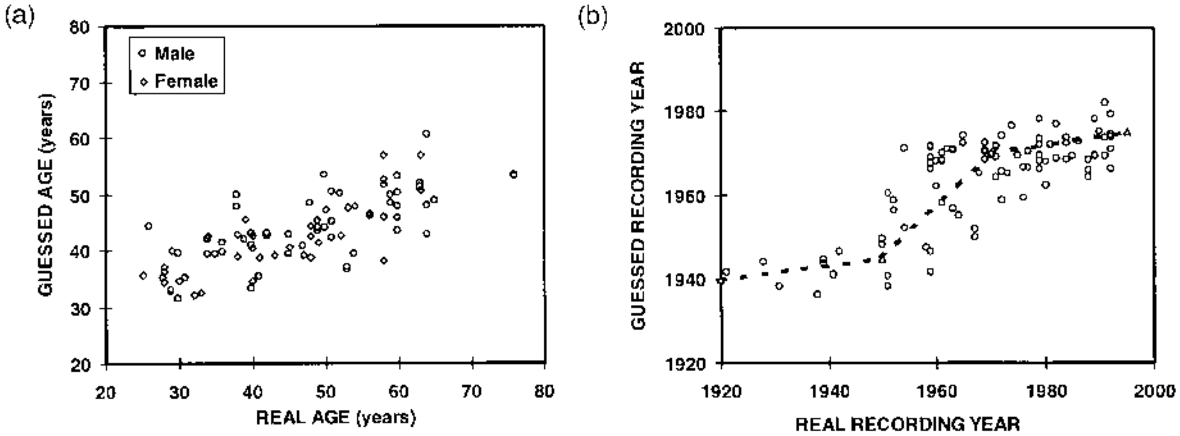

Figure 1. (a) Expert listeners' ratings of the age of twenty classically trained singers, representing the categorisations of soprano, alto, tenor, baritone and bass, recorded over a period of 40 years. (b) Expert listeners' ratings of the year of recording for the same material as shown in (a). The dashed line represents an approximation of the overall trend (from Sundberg *et al.* 1998).

main cause; rather, the voice timbre was the main problem. An acoustic analysis of Burns' voice therefore promised to be a rewarding task.

The timbral characteristics of a voice derive from two sources (Sundberg 1987). One is the *voice source*, i.e. the pulsating transglottal airflow. This airflow signal is determined by three main physiological parameters. One is the overpressure of air in the lungs, or the lung pressure, which is controlled by the respiratory system. The second is the dimensions of the vocal folds which are determined by the muscles that stretch and stiffen the vocal folds. The third is the degree of vocal fold adduction, which reflects the degree to which the vocal folds are pressed against each other. Adduction is controlled by a group of laryngeal muscles.

The acoustic consequences of changes in these physiological parameters are reasonably well understood. Lung pressure determines vocal loudness, which is closely related to the sound level. A discontinuity in the transglottal airflow waveform occurs when the vocal folds close the glottis. The lung pressure controls the sharpness of this waveform discontinuity, and so determines the amplitudes of the lower overtones which, in turn, decide the overall SPL of the sound. The length and stiffness of the vocal folds determine their frequency of vibration, i.e. the fundamental frequency of the sound. Glottal adduction determines the mode of phonation, a phonatory dimension with the extremes of voiced whisper and pressed phonation corresponding to minimum and maximum adduction, respectively. It determines the amplitude of the translottal air pulses, such that a high degree of glottal adduction results in short air pulses of low amplitude. Such air pulses produce a weak fundamental.

Long-term-average spectrum (LTAS) analysis of singers of the same voice classification can reveal who has the weaker fundamental, and thus who sings with more glottal adduction. The comparison is particularly reliable if the singers sing the same melodic sequence at the same pitches. Figure 2 shows such LTAS analyses of Burns and another, unarguably spotless voice, that of Nicolai Gedda. Burns' fundamental is about 6 dB weaker than that of Gedda. It seems likely that this is not due to recording technology, but rather to the singing technique. Thus, the LTAS suggests that Burns was using a firmer glottal adduction than Gedda, i.e. his phonation was more pressed. This would account for part of the great timbral differences between these two voices.

The other factor of major importance to the voice timbre is the constellation of vocal tract resonances. These resonance or *formant* frequencies decide vowel as well as personal voice quality. Our sensitivity to formant frequencies is great; the difference limen in formant frequency is about 3%, or about 50 cent (Flanagan 1955, Nord and Sventelius 1979). This value is relevant both in sampler technology and in audio processing; human voices need to be sampled at each semitone along the scale in a sampler, and the possibilities to change sampling rate without affecting voice timbre are limited.

The formant frequencies are controlled by means of the vocal tract shape and thus by the positioning of the articulatory organs such as the lip and jaw openings, the tongue shape and the larynx position. In male singers and altos, who are trained according to the western classical opera tradition, an important acoustic characteristic is the singer's formant. This is a spectrum envelope peak near 3 kHz (Sundberg 1987). It is produced mainly by a clustering of the third, fourth and fifth formants; in other words, by resonance rather than vocal effort. It is perceptually important, as it helps the singer's voice to be heard when the orchestral accompaniment is loud. The background to this is that the sound of the orchestra is comparatively weak in the frequency range of the singer's formant. The two LTAS in figure 2 reveal a great difference between these two singers with regard to their formants; in Gedda's LTAS the level near 3kHz is about 20 dB higher than in Burns' LTAS. This would

https://doi.org/10.1017/S1355771800003058 Published online by Cambridge University Press

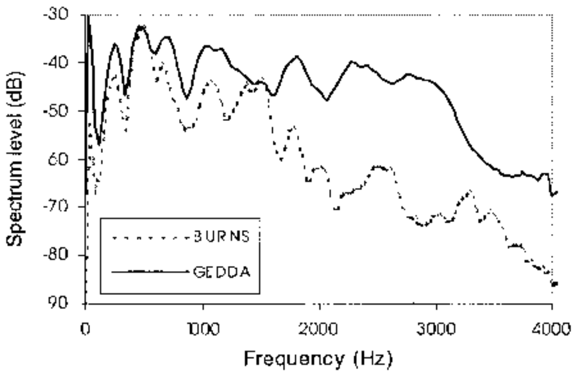

Figure 2. LTAS analyses of about 13 s of solo singing from Charles Gounod's opera Faust, performed by T. Burns and N. Gedda (CD recordings: GD 61175, BMG Classics, *The Glory (????) of the Human Voice* and CMS 7 69983 2, EMI Classics, *Gounod Faust*).

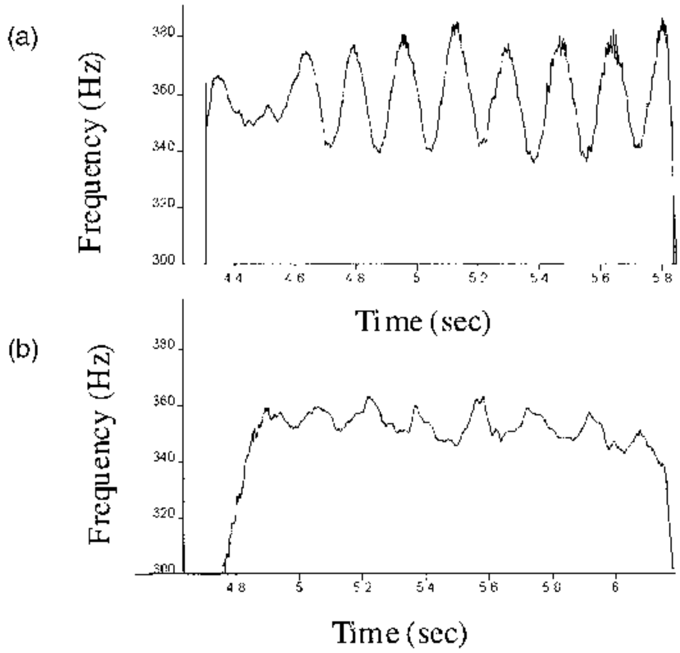

Figure 3. Fundamental frequency analysis of (a) N. Gedda's and (b) T. Burns' renderings of the pitch g4 taken from the recordings specified in figure 2.

account for much of the timbral differences between these two singers.

Figure 3 compares the two voices' renderings of the pitch G4. The mean fundamental frequency is 31 cent flatter in Burns, so he is clearly out of tune. The mean peak-to-peak amplitude of the vibrato undulations is 90 cent in Gedda and 28 cent in Burns, and the variability of this amplitude is much greater in Burns, 48% rather than 20% of the mean amplitude. The regularity of the vibrato frequency, measured as the standard deviation of vibrato frequency data for each half vibrato period amounts to 0.4 Hz in Gedda and 2 Hz in Burns. Thus, the two voices differ with respect to the mean fundamental frequency, the mean extent and the regularity of the vibrato.

Synthesis experiments confirm that specific characteristics of both glottal adduction and vocal tract shape contribute significantly to the unpleasantness of Burns' singing voice. This does not answer the question of why his voice possesses this striking, primitive ugliness, but it is tempting to speculate. Ease of production seems to contribute importantly to an aesthetical quality. Burns' lack of a singer's formant implies that he attains audibility simply by singing loudly, rather than by using the effortlessly available potential of vocal tract resonance. Burns' relatively low amplitude of the fundamental would contribute to the impression of pressed phonation, i.e. a high degree of vocal effort. There are reasons to assume also that his narrow vibrato extent contributes to the same impression (Sundberg and Askenfelt 1983);

https://doi.org/10.1017/S1355771800003058 Published online by Cambridge University Press

listeners tend to perceive a reduced vibrato extent as a sign of pressed phonation. Thus, the unpleasantness of his voice may be explained by an ensemble of signs of a forced voice production.

How does this translate into music technology and audio processing? These analyses of a good and a bad singing voice seem instructive. Our daily experiences of voice quality are likely to bias our timbre perception. The amplitude of the fundamental is also likely to be a particularly relevant factor in other instruments. For example, the long strings and large sound board dimensions of a grand piano produce louder low-frequency partials than short, low-quality upright pianos. Likewise, high-quality double basses seem to produce louder low frequency components than low-quality double basses.

The reason for our preference for a strong fundamental may not merely be brain-washing caused by our acquaintance with human voices. At most pitches (fundamental frequency >100 Hz, approximately), the fundamental resides in a critical band of its own, surrounded by silent critical bands. This suggests that it contributes importantly to timbre perception. The commonly used linear frequency scale in line spectrograms may tend to seduce us into underestimating the importance of the fundamental; in any event, a perceptually realistic appreciation of a spectrum is likely to be promoted by auditorily realistic spectrum analyses.

A well-controlled vibrato may represent another important feature of musical sounds. In singing, the vibrato frequency is mostly a personal constant, which the singer cannot control, although the frequency tends to increase somewhat toward the ending of tones (Prame 1994). The amplitude of the fundamental frequency undulations, on the other hand, tends to depend on the loudness of the tone. Also, it seems typically to be great in agitated pieces and small and smooth in pieces with a calm ambience. It might also be that the waveform of the vibrato undulations is varied in a meaningful way. This is in clear contrast to the typical vibrato available in many synthesizers, where the amplitude is constant and the waveform is mostly purely sinusoidal. An important task of music acoustics is to find out the principles according to which the vibrato parameters are varied in real performances.

The singer's formant deserves some more comment. First, the singer's formant resides in the frequency range of the third formant, which is crucial to the identification of consonants. It is therefore likely to contribute to the intelligibility of the text. It is also interesting that a filter which raises the level in the frequency range of the singer's formant is referred to as a 'presence filter'. Second, the singer's formant seems optimally adapted to the LTAS of the typical orchestral accompaniment in operas, as mentioned. One would then expect that no singer's formant should develop in repertoires where the spectral properties of the accompaniment are different

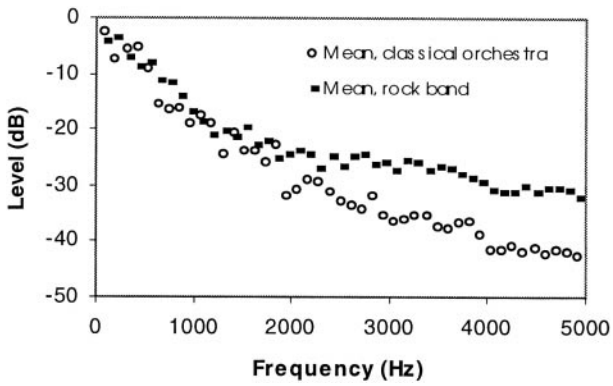

Figure 4. LTAS of accompaniments typically used in rock music bands and of a traditional operatic orchestra (after D. Zangger Borch, unpublished Master Thesis, 1998, in Swedish).

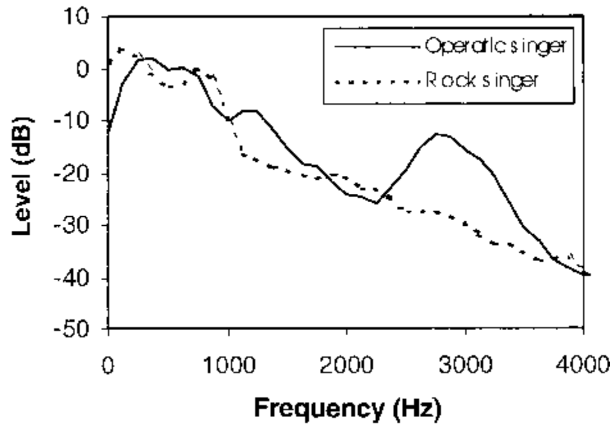

Figure 5. LTAS of a rock music and an opera singer performing the same, 20 s long, music excerpt.

from those of the western operatic tradition. In figure 4, an LTAS of accompaniments typically used in rock music is compared with that of the traditional operatic orchestra (Zangger Borch 1997). It can be seen that the spectrum level near 3 kHz is about 10 dB lower in the operatic orchestra. Thus, from the point of view of audibility, a singer's formant would be more profitable for an opera singer; as illustrated in figure 5, there is a considerable difference in the LTAS levels near 3 kHz between an opera singer and a rock singer. There are no signs of a singer's formant in the rock singer's voice.

4. PERFORMANCE

Most musical performances deviate significantly from deadpan, as mentioned previously. Our research in this area has resulted in a generative grammar that automatically converts input note files into sounding performances on a synthesizer (Friberg 1995). The grammar consists of some twenty context-dependent rules that introduce micropauses, lengthenings and shortenings of tone durations, and long- and short-term increases and decreases of sound level.

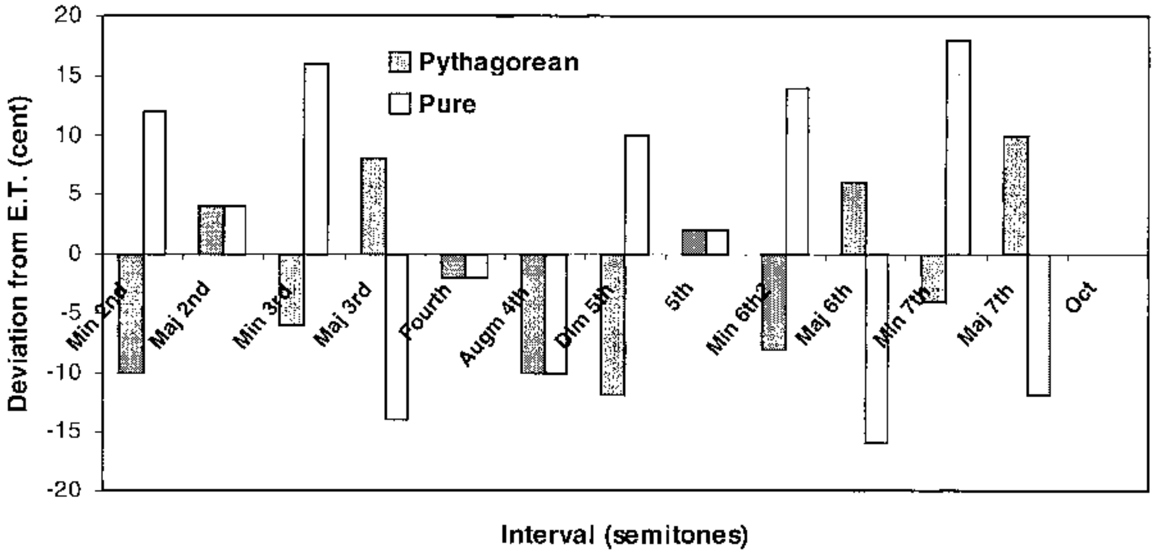

Figure 6. Difference between Pythagorean and just tuning of some intervals, given as deviations from the equal tempered tuning.

The rules can be divided into three groups according to their apparent function in music communication. One group, the *ensemble* rules, achieves agreement in timing and tuning in ensembles. Another group, the *differentiation* rules, enhances the differences between pitch and duration categories. A third group, the *grouping* rules, introduces markings of structural elements such as melodic gestures/motifs, subphrases and phrases. Some examples of performance rules will next be presented.

The tuning of instruments in ensemble playing is an area of confusion. The equal tempered tuning, where all minor second intervals of the scale have the fundamental frequency ratio of $1{:}2^{(1/12)}$, is normally used in traditional keyboard instruments, such as the organ and the harpsichord, and basically also in pianos. Yet, just intonation, where the tones of the diatonic scale are derived from a just tuning of the three main harmonies, i.e. tonic, dominant and subdominant, seems to exert a somewhat romantic attraction on music theorists. The advantage of just tuning is that it eliminates beats in consonant chords played on instruments with exactly harmonic partials, since the frequencies of common partials then agree exactly. Indeed, an invariably just tuning is used as a selling point for some synthesizers.

Analyses of music performance, however, have revealed that musicians, who are free to decide on the tuning of their instruments, do not stick to just, but rather to the Pythagorean tuning. Here, the frequencies of the diatonic scale tones are derived from a chain of pure fifths. This produces an important deviation from just tuning, in that minor seconds, thirds and sixths are about 20 cent narrower than in just tuning, while the major versions of these intervals are wider than in just (figure 6). This appears to add a desirable quality to intonation of melodic intervals, i.e. between successive tones. The reason can be assumed to be the differentiation of categories. Minor and major melodic intervals would represent an important categorisation in music listening. The application of Pythagorean tuning of melodic intervals enhances the difference between these categories by making minor intervals narrow and major intervals wide as compared with both equal tempered and just tuning.

The disadvantage with the Pythagorean tuning is that it gives rise to beats in consonant chords, as long as tones with harmonic partials are used. The strategy applied in our performance grammar is a combination of the Pythagorean and just tunings (Sundberg, Friberg and Frydén 1989). For melodic intervals, the Pythagorean tuning is applied, such that the minor versions of intervals become narrow. For dyads and chords sustained long enough to produce salient beats, the tuning is slowly changed during about a second until it eventually reaches the just version of the intervals.

The marking of structural elements such as melodic gestures, subphrases and phrases, seems a basic principle in music performance. Todd (1985) found that subphrases and phrases were started with an *accellerando* from a slow tempo and terminated with a *rallentando*. This pattern has later also been found in sung performances (Sundberg 1998). Indeed, even when a professional singer was asked to perform deadpan versions of some music excerpts, he still tended to introduce these phrase markings (figure 7). These timing deviations from deadpan can be reproduced rather accurately by our phrase rule.

Deviations much smaller than those shown in figure 7 are often found in performances, and our sensitivity for departures from an exact isochrony is very high; deviations as small as about 10 ms can be noted for tone durations between 100 and 250 ms, approximately

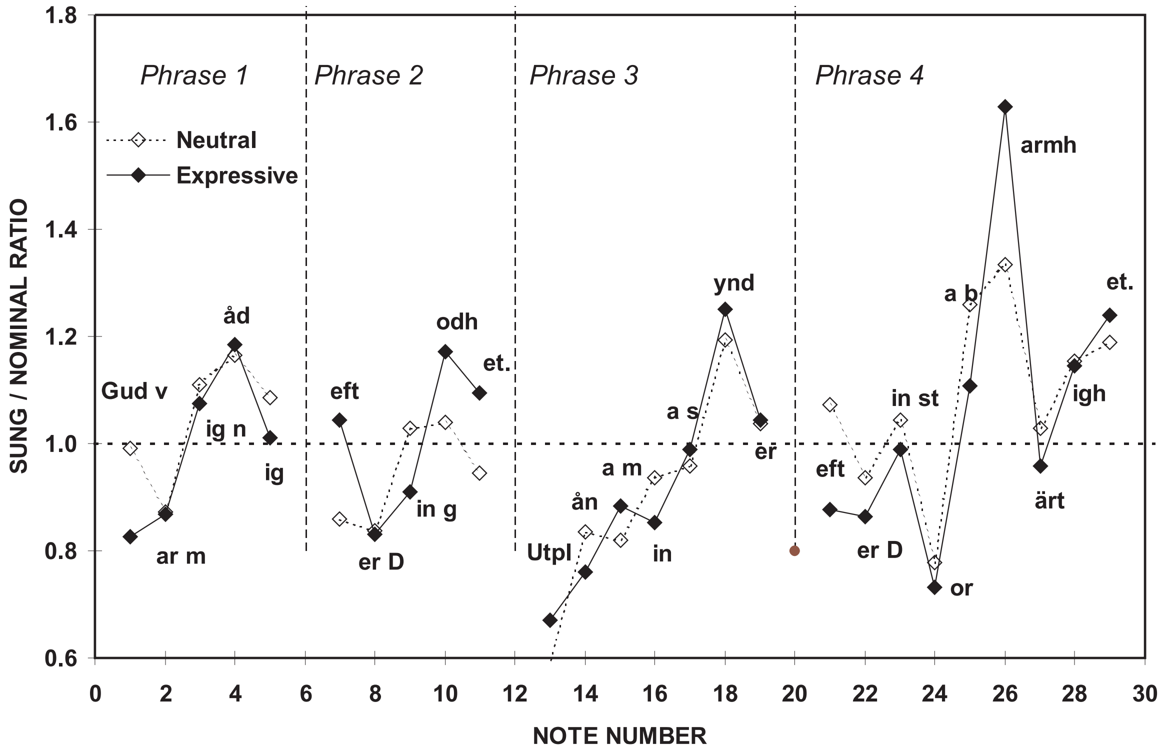

Figure 7. Professional baritone singer's deviations from nominal tone durations in a typical and in a deliberately unexpressive rendering (solid and dashed curves, respectively) of the first bars of aria no. 18 from F. Mendelsohn's oratorio Paulus, op. 36 (Sundberg, Iwarsson and Hagegård 1995).

(Friberg and Sundberg 1995). Thus, timing seems to be an exceedingly productive dimension of music performance in the sense that much information can be conveyed by minute departures from nominal tone durations. Therefore, sequencer programs that simply reproduce nominal durations do not seem promising; the lack of timing deviations tend to produce boring, machine-like performances. A generative grammar that produces meaningful deviations in timing seems a necessary complement to worthwhile sequencer programs.

These examples demonstrate some aspects that seem important to music performance and hence also to the construction of synthesizers and to audio processing. The equal tempered scale, generally assumed to represent ideal intonation, is in fact a poor model. Musicians playing instruments with variable intonation tend to tune the tones depending on the musical context. The variations are mostly minute, but certainly relevant to a musically trained listener (Sundberg, Prame and Iwarsson 1996). A musically attractive synthesizer would need to offer the possibility of context-dependent fine tuning. In the case of automatic performance of music files, this can be achieved by implementing in the sequencers tuning recipes of the type described above. In the case of synthesizers, performer-controlled intonation seems indicated. Likewise, musicians' deviations from nominal tone durations are certainly relevant to the musical quality of the performance, even if they are sometimes minute. Editing of tuning and timing in a music recording would be greatly helped by acquaintance with the theory of music performance.

5. OUTLOOK

The title of this article raised the question of whether music technology and audio processing are facing a *rallentando* or an *accelerando* into the new millennium. The answer seems to depend on two major factors: technology and knowledge.

With respect to technology, there seems to be no reason to assume that the present gallop will slow down. Synthesizer technology is charging ahead apace with the proliferation of low-cost, generally available computing power, even allowing the implementation of complete numerical models of traditional instruments. The Internet offers facilities to everyone for playing music. Digital sound processing has made handling of recorded sound easy, and the possibility to do almost anything you can think of is available today.

Technical development has to be supported by an expansion of knowledge, to ensure that the result will not be an unmusical technology, likely to eventually bore music listeners. The present rate of development of new knowledge about music can perhaps best be described as a trot compared to that of music technology.

The questions HOW? and WHY? keep producing relevant information, however, and understanding of timbral beauty and ugliness is under development. Music performance has been the target of a massive research effort over the last twenty years, revealing a completely new view of, for example, tuning and timing practice in high-quality performances.

An important goal must be to implement all this knowledge in music technology and audio processing. Future construction of synthesizers will profit from the understanding of timbral beauty and ugliness, as well as the tuning practice in top-quality music performance. Skilled musicians' substantial departures from nominal durations in the score will help music programmers realise that deadpan performance is a musically dead horse. They will feel the urge to implement performance grammars in future music programs, which users will be able to manipulate so as to experience participation in the process of creating satisfying performances.

Thus, many of the fruits of music research are yet to be used in music technology, and the implementation of this knowledge is likely to stimulate the demand for music technology. Therefore, there are good reasons to assume that music technology and audio processing are now facing an *acclererando* into the new millennium.

ACKNOWLEDGEMENTS

This is a revised version of the author's keynote speech at the 25th Euromicro Conference in Milan, Italy, September 1999.

REFERENCES

Flanagan, J. 1955. A difference limen for vowel formant frequency. *Journal of the Acoustical Society of America* **27**: 613–17.

Friberg, A. 1995. *A Quantitative Rule System for Musical Expression*. Dissertation, KTH.

Friberg, A., and Sundberg, J. 1995. Time discrimination in a monotonic, isochronous sequence. *Journal of the Acoustical Society of America* **98**: 2,524–31.

Nord, L., and Sventelius, E. 1979. Analysis and prediction of difference limen data for formant frequencies. *STL-QPSR* **3–4**/1979: 60–72.

Prame, E. 1994. Measurements of the vibrato rate of ten singers. *Journal of the Acoustical Society of America* **9**: 1,979–84.

Sundberg, J. 1987. *The Science of the Singing Voice*. Dekalb, IL: Northern University Press.

Sundberg, J. 1993. How can music be expressive? *Speech Communication* **13**: 239–53.

Sundberg, J. 1998. Expressivity in singing. A review of some recent investigations. *Logopedics, Phoniatrics, Vocology* **23**: 121–7.

Sundberg, J., and Askenfelt, A. 1983. Larynx height and voice source: a relationship? In J. Abbs and D. Bless (eds.) *Voice Physiology*, pp. 307–16. Houston, TX: Collegehill.

Sundberg, J., Friberg, A., and Frydén, L. 1989. Rules for automated performance of ensemble music. *Contemporary Music Review* **3**: 89–109.

Sundberg, J., Iwarsson, J., and Hagegård, H. 1995. A singer's expression of emotions in sung performance. In O. Fujimura and M. Hirano (eds.) *Vocal Fold Physiology. Voice Quality Control*, pp. 217–31. San Diego: Singular Publishing Group Inc.

Sundberg, J., Niska-Thörnvik, M., and Söderström, A.-M. 1998. Age and voice quality in professional singers. *Log. Phon. Vocol*. **23**: 169–76.

Sundberg, J., Prame, E., and Iwarsson, J. 1996. Replicability and accuracy of pitch patterns in professional singers. In P. J. Davis and N. H. Fletcher (eds.) *Vocal Fold Physiology. Controlling Complexity and Chaos*, pp. 291–306. San Diego, CA: Singular Publishing Group.

Ternström, S. 1981. A synthesizer of bell and chime sounds – ISABEL. *STL-QPSR* **2–3**/1981: 17–22.

Todd, N. 1985. A model of expressive timing in tonal music. *Music Perception* **3**: 33–58.

Zangger Borch, D. 1997. Unpublished Master Thesis (in Swedish), Royal Conservatory of Music, Stockholm.