MUSIC ACOUSTICS

The music acoustics group is presently directed by a group of senior researchers, with professor emeritus Johan Sundberg as the gray eminence. (from left Johan Sundberg, Roberto Bresin, Sten Ternström, Anders Friberg, Erik Jansson and Anders Askenfelt). Photo by C Dunger.

The Music Acoustics group is the second largest at the department. During 2001, five senior scientists, five research students, eight international guests and three project workers, in all 25 persons, were working in the group. In July 2001, professor Johan Sundberg retired. Following the current policy of KTH, the succession will be implemented at some later date by promotion of one of the senior lecturers. The music acoustics group is currently directed by an executive committee consisting of the five senior group members.

Last year the music acoustics group was appointed to be a *Marie Curie Training Site* during 2001 – 2004, which allows us to receive visiting doctoral students for a period of up to one year. This is a gratifying recognition of the activities within the group, which also will help us to continue and extend future co-operation with our colleagues in the European countries.

Voice

For many people the voice is the work tool number one, day after day. This applies not only to singers and actors, who use their voices to carry an artistic message across to the listeners. Teachers of children in all ages, salesmen, consultants and many more speak almost continuously in their work, often in busy environments. While singers and actors have trained their voices, others have not and may experience voice dysfunctions. Reliable voice diagnostics based on easy-to-perform acoustical analyses is an important medical research area in which the group has a long-standing co-operation with the Department of Phoniatrics and Logopedics at Huddinge University Hospital.

A deeper understanding of the phonatory mechanism and the vibrations of the glottal folds is essential in all studies of speech and singing. Within this research area we are continuously exploring new methods for measurements and modelling. An important field of application is the singing voice, trying to explain professional singers´ phonatory and articulatory techniques.

Effects on voice habits of noisy work environments

Workplace noise is a probable cause of voice problems in several professions. In a quiet clinic, it is not easy to find out what the patient really does with her voice when confronted with the real noise. Recordings on location precludes certain types of analysis, since they will be contaminated by the offending noise itself. In a new project, we re-create the noisy environment in the laboratory over hi-fi loudspeakers. The subject tries to make herself heard over the simulated noise, and the voice is recorded together with the noise. This recording can be so noisy as to be unintelligible. Current channel estimation techniques allow us to cancel out the noise and recover the voice signal. Up to 40 dB of noise rejection can be obtained. The noise signal can be any appropriate sound, such as hi-fi recordings of actual workplace noise. With this setup, the clinician can provoke a realistic sample of voice abuse, acquire a recording that is clean enough for standard acoustic analyses, play it back to the patient, and explore alternative vocal strategies for workplaces where the noise cannot be abated. Measurements of the phonatory habits of normal speakers in loud environments will supplement the analyses of voice patients.

Measurement of total speaking time

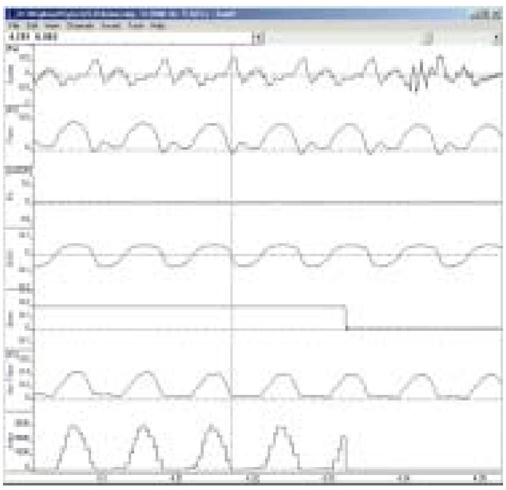
Voice load is a function of speaking level and time. To gain data on the voice usage during work we are developing a method for separating a speaker’s voice from simultaneous environmental sound. The method processes the signals from two binaurally mounted microphones, the signals from which are nearly identical only for the speaker’s voice. The method was test ed on a group of pre-school teachers who were recorded during work hours, using portable tape recorders. The teacher’s voice was automatically separated from the environmental sound and the total voicing time, mean fundamental frequency, speaking level, and the level of the environmental sound were calculated. The method allows a complete, automized documentation of voice usage in work and everyday live, including the sound environment which the subjects have to compete with.

Resonance frequency of the vocal folds

Fundamental frequency in phonation is influenced by the resonance frequency of the vocal folds. Exploring the potential of an old idea, we have tried to excite the vocal folds by means of a sine tone introduced into the vocal tract, and record the response of the folds to this acoustic excitation by high-speed imaging. As the tension of the vocal folds and their corresponding resonance frequency varies depending on the intended phonation frequency, the subject needs to phonate at a given pitch and then slightly abduct the folds while attempting to keep vocal fold properties constant. Although this may seem quite difficult for an untrained subject, the results were in good agreement with expectations and previous studies. The method will be used in further studies of the vocal folds in voice patients and healthy subjects.

Vocal fold vibration and pulsating transglottal airflow

The relationship between glottal area and transglottal airflow has been difficult to analyze in detail. We have studied this aspect of the phonatory mechanism in co-operation with Huddinge University Hospital, using a combination of high-speed digital imaging of the vibrating vocal folds and inverse filtering of the associated airflow. Glottal area was measured by a custom-made computer program. The results confirmed expected relationships between flow and area, even though the relationship varied considerably depending of the mode of phonation.



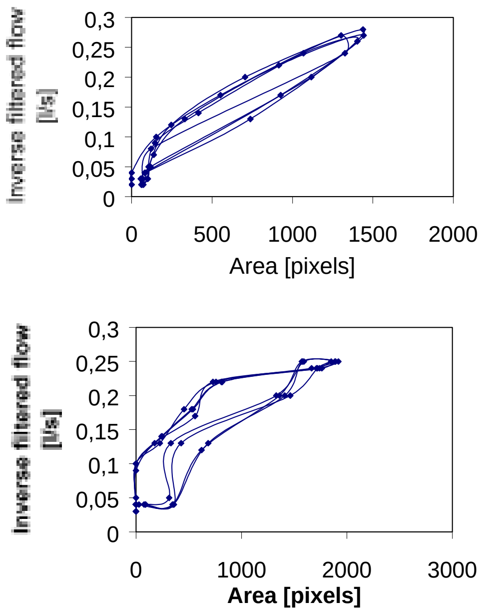
Signals recorded during the high-speed imaging experiment. From top: audio, flow, oral pressure (irrelevant durin vowel phonation), EGG, synchronization pulse, inverse filtered flow, glottal area. The flow and audio signals have been shifted in time to correct for the delay caused by the speed of sound propagation.

Velopharyngeal opening in singers

Many operatic singers sing the vowel [a] with a velopharyngeal opening (an opening between the vocal tract and nasal cavity). The advantage of such an articulatory technique was studied using CT scan measurements of a baritone singing the vowel [a]. By acoustic measurements on a three-dimensional epoxy model of the vocal and nasal tract, made according to the CT scan measurements, it was showed that the velopharyngeal opening attenuates the first formant of the vowel [a]. This may increase the relative prominence of the singer's formant, giving a more brilliant timbre.

Broadway and classical vocal style

The vocal differences between the typical style of Broadway performances (musicals)



Relationship between transglottal airflow, derived by inverse filtering, and relative glottal area, recorded by high-speed imaging, for neutral (above) and pressed-/hyperfunctional (below) phonation.

and operatic styles of singing have been studied in co-operation with Vanderbilt Voice Center, Vanderbilt University, Nashville, USA. An artist, professionally experienced in performing in both styles, sang the same material in both styles. Long-time-average-spectra analyses revealed that, in Broadway style, the fundamental is less prominent than higher partials, indicating higher subglottal pressures and a firmer adduction. By and large, the Broadway style appeared more speech-like than the operatic style.

Glottal source modeling

In co-operation with University of Padova a physically-informed model of the glottal source that uses a combination of linear and non-linear elements is being studied. The key feature of the model is that the non-linear element is self-configured by training against real flow waveform data. A detailed modelling of the vocal fold mechanics is thus not needed. This work is conducted within the framework of the Marie Curie Training Site program.

Pevoc IV

An important undertaking in the voice area during 2001 was the organization of the 4th Pan-European Voice Conference (PEVOC IV), which attracted more than 300 participants. The conference focused on four themes:

- Basic biological aspects – Applications to voice disorders
- Voice source - Acoustics and imaging
- Assessment & efficacy of voice treatment
- Singing and professional speaking

Musical Instruments

The physics of the tone generation in the traditional stringed instruments like violin and piano has been a long-standing interest in our group, as well as the relations between tone quality and instrument design. In parallel, our interest in studying musicians' control of their instruments in playing has grown, and has now expanded into a research theme of its own. One reason is the many applications in the design of synthesizers and the associated man-machine interfaces, where the keyboard has been the universal but, in many cases, not the optimal solution. In computer-generated music, too, the interest for human-like control is growing.

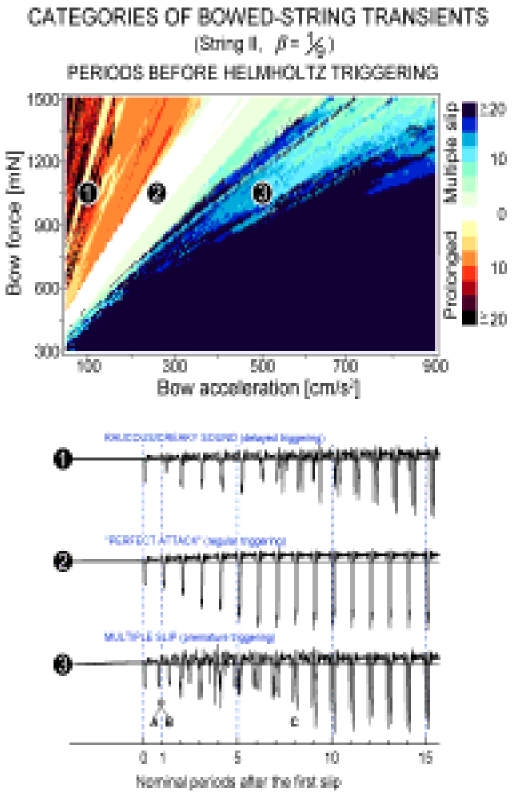
Violins

The research on the violin itself has concentrated on the relative importance of the bridge and top plate. A highly desirable property of a violin is a peak in the sound spectrum at about 2.5 kHz, adding to the "brilliance." This peak may be boosted by tuning (cutting) the bridge to a fundamental resonance in this frequency range. The basic conditions for a 2.5 kHz-peak in the radiated sound must, however, be present in the instrument itself. The crucial design parameter seems to be the stiffness in the cross direction of the top plate at the position of the bridge

Violin bows

The individual properties of violin bows are very important to the professional player, who selects a bow with great care. Among other things, the bow is claimed to influence the timbre of the particular violin to a large extent. Here, too, a cross-modal interference cannot be excluded. The violinist is "in the loop," controlling all parameters in the bowing gestures, including the bow velocity and bow force against the string. The auditory impression may well be influenced by the feel of the bow via the kinesthetic forces in the bowing hand. Even more important, the player is well aware of all compensating actions which are made in order to reach the desired sound volume and timbre, as well as "clean", scratch-free attacks of the notes. In listening tests, it was shown that professional violinists are not able to discriminate between high-quality and medium-quality bows by listening only. This contrasted against their judgments in playing, which were much more discriminating. Like the pianists, the violinists seem to get an exaggerated view of the influence on timbre compared to the listeners.

Helmholtz motion



Computer simulations of transients in a bowed string ("attacks"). Each pixel represents a certain combination of bow acceleration and bow force. The color code reflects the duration of the aperiodic part of the transient. Musically acceptable attacks are restricted to the areas in white and light colors.

The conditions for development of the normal periodic violin tone, the so-called Helmholtz motion, is being analyzed in a doctoral thesis work. The bowed string has been studied for 150 years, and the steady-state behavior is rather well understood, including the effects of string stiffness and finite width of the bow hair among other things. The transient process at the start of a note (“the attack”), however, has not been thoroughly addressed. During the attack, the string must quickly be set in motion, ideally with a periodic motion from the first period. It turns out that the main control parameters are the acceleration of the bow and force between bow and string (bow force).

Parallel simulations, including a wealth of bowing parameter combinations, show that the margins are very narrow for obtaining, and maintaining, a periodic motion during the first periods. Nevertheless, it has been shown that professional players have little difficulty in achieving long series of “perfect” attacks, if desired.

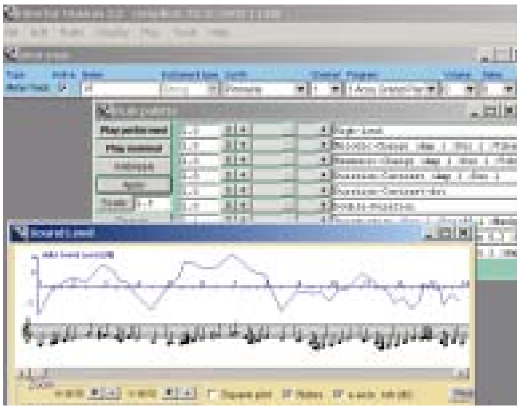
The pianist’s touch

The classical controversy between pianists and physicists about the pianists “touch,” meaning the possibilities for the pianist to influence the timbre or character of individual piano tones by the way the key is depressed, has been revisited. Literature reviews, interviews of pianists over Internet, and laboratory experiments support the notion that pianists to a large extent are subject to an illusion. The auditory perception of the tone seems to be “colored” by the kinesthetic feed-back in the fingers from the keyboard and action, and the pianists are fooled by their own senses into believing that the tone can be significantly influenced by touch. An action of poor mechanical quality may be perceived as a deficiency in timbre, caused by the hammer, strings, or soundboard. Such a cross-modal interference is well known to influence the perception of sound in other situations. For example, the perceived interior noise level in cars is dependent on the vibration level in the seats and on the view through the wind screen. In our case, the pianists’ impression of the tone is strongly touch-dependent, but the acoustical process much less so. The listeners, who only have their ears to rely on, would perceive a very limited influence of the pianists’ touch, compared to the performers themselves.

Performance

Director Musices

Studies of musical performance is a core area in the research in the music acoustics group. Starting back in the 70´s, a generative grammar for the performance of music based on common music notation has been developed, implemented in the program *Director Musices*. Since a few years it has been available on the Internet, attracting a large interest from the very beginning. *Director Musices* automatically converts music files in MIDI to sounding performances possessing musical expressivity. Lately, options for emotional coloring of the performances have been introduced. The applications are many. Today when music files in MIDI format are distributed and sold over Internet *en masse*, tools for musically acceptable performances of downloaded music are needed more than ever. A mechanical reproduction of a melody in common music notation usually defines a "standard of boringness." The *Director Musices* will cure this syndrome, giving a musically decent performance.



Another application of the rules for musical performance with high commercial potential is in ring tones of mobile phones. Besides from being able to select your personal ring tone melody, the mood in which it is performed may be made dependent on who is calling. A call from your wife or girl friend could sound in a quite different mood compared to that when your boss is calling. In addition to numerous applications of the performance rules, the *Director Musices* program serves as a workhorse in a number of research projects.

EU projects

Several EU projects with partners in different European countries, in particular Italy, relate to musical expression. In *MEGA* ("Multisensory Expressive Gesture Control"), a co-operation with Universitá di Padova and others, the aim is to extract, synthesize and map expressive features in dance, musicians' gestures, and music performance. In one part of the project, the expressive information in the motion of percussion players´ hands, arms, and body is being studied. As an example, preliminary models indicate that it is possible to use automatic cue extractions of the motions of the upper arms of a marimba player in order to catch the intended mood of the performance (angry, solemn, happy and scared). The result tells us something about the many parallel channels emotional content in music is transferred from the performer to the "listener". Even when deprived of the information carried by the sound, face expression, and the movements of the hands and mallets, subjects were still able to identify the intended mood of the performances.

Another subtask in MEGA is to distill the acoustical characteristics that carry the information on the emotional expression. A number of parameters like tempo, lengthenings and shortenings of note durations, and variation in dynamic level contribute in this information transfer. A method for automatic extraction of key parameters (cues) setting the mood of the performance has been developed, applicable on various music styles and types of musicians.

A somewhat similar approach is taken in the Swedish-supported project *Feel-Me* ("Feedback Learning in Musical Performance"), conducted in co-operation with the Dept. of Psychology, Uppsala University. The aim here is to apply the results from performance research in teaching students to play expressively. At basic levels of music teaching, the playing technique, including all its practical and motoric difficulties, is often put in focus at the expense of the emotional expression. The idea in *Feel-Me* is to extract some of the most important parameters of the performance carrying emotional content, and display a summarized measure of how pronounced the student performs a particular mood, for example, “happiness”. A new and empirically-based approach to learning emotional communication, coined cognitive feedback, will be developed and implemented in user-friendly software. The project involves an interdisciplinary collaboration between psychologists, technicians, music teachers, and musicians.

In *SOb* (“The Sounding Object”) the aim is to develop sound models which can be connected to various objects in a virtual world. As an example, results from our previous performance research related to human gait has been applied. By applying rules originally aimed for describing the performance of *staccato*, *legato*, and *ritardando* in music, it was shown possible to control the sounds of footsteps in such a way that the listeners easily can discriminate between walking and running.

Mosart (“Music Orchestration Systems in Algorithmic Research and Technology”) is a EU training network focusing on music performance issues. A number of guest researchers have visited us from European universities, working on emotional and gestural content in music performance, as well as modeling violin vibrato.

“Radio Baton”

A particularly interesting project with a visionary flavor is an application in which the performance rules of *Director Musices* are combined with “The Radio Baton.” The baton system is a manually controlled sequencer, in its appearance much like a conductor´s ordinary baton, developed by one of the pioneers in computer music, Max Mathews at Center for Computer Research in Music, Stanford University. By conducting with the radio baton in a cube-shaped space over a sensor area, individual notes, bars, or whole phrases are triggered one after the other. Once an event has been triggered, *Director Musices* organizes the performance within the event, including musical expression and emotional coloring. The conductor *in spe* can thus chose the level at which she wants to interfere at different points in the music, and control tempo, *rubato* etc in a manner much like an ordinary conductor. Apart from pedagogical purposes in the training of conductors and other musicians, the application seems to have potential for orchestral accompaniment of soloists in performances of contemporary music.

Other projects in the area of music performance include the building of a rule-based generative grammar for the performance of vibrato in string music, studies of the *staccato-legato* dimension in piano performances, and motion strategies used by drummers for performing accented notes interleaved in a steady rhythm.

Filename:	MUS4.doc
Directory:	I:\annual\Annual2001
Template:	J:\tmh.kth.se\tmh\proj\common\template\tmh\Normal.dot
Title:	
Subject:	
Author:	Personal
Keywords:	
Comments:	
Creation Date:	2002-04-16 10:47
Change Number:	19
Last Saved On:	2002-06-11 15:24
Last Saved By:	Personal
Total Editing Time:	285 Minutes
Last Printed On:	2002-06-12 10:41
As of Last Complete Printing	
Number of Pages:	8
Number of Words:	2 769 (approx.)
Number of Characters:	15 787 (approx.)