# EXPERIMENTS ON PITCH PERCEPTION: IMPLICATIONS FOR MUSIC AND OTHER PROCESSES

Adrianus J. M. HOUTSMA

US Army Aeromedical Research Laboratory
Building 6901, Fort Rucker, AL 36362, USA

This paper presents a review of the development of our understanding of pitch perception during the past three decades, in which the author was deeply involved. Some specific applications are discussed that are for the most part in the domain of music perception and show where the interest in pitch perception has led us today.

**Keywords:** pitch perception, low- and high-order harmonics, harmonic intervals, polyphony, homophony.

## 1. Prologue

This paper does not focus on any single specific topic. Instead, I have attempted to describe a few elements of my research career and have tried to show how they were often inspired by and interacted with ideas and thoughts colleagues and students. It involves work that was done while at the Research Laboratory of Electronics at MIT, at the Institute for Perception Research at the Technical University of Eindhoven in The Netherlands and, currently, at the US Army Aeromedical Research Laboratory at Fort Rucker in the USA.

## 2. Pitch of complex tones

During my PhD study at M.I.T. I became very interested in a pitch phenomenon that was described in the literature as the "case of the missing fundamental". It is something that we all have experienced when listening to speech or music played through a small loudspeaker ( e.g., a telephone, a transistor radio or a laptop computer) that transmits only high frequencies. For low-pitch musical instruments like double bass, cello, bassoon, or a male singing voice, only harmonics will be heard since the loudspeaker cannot transmit the fundamental frequency. This decreases the sound quality, but generally does not confuse the listener in hearing what notes or melodies are being played.

Somehow, our ears and brain are able to reconstruct the correct pitch of a note, which is directly correlated with the fundamental frequency of the sound, even if there is no physical energy at that frequency.

This pitch phenomenon had been known for quite some time. Ohm and Seebeck had animated discussions in the Annalen für Physik und Chemie of the 1840s on whether it really existed and what its cause might be. HELMHOLTZ explained it as a result of signal distortion in the middle ear, causing a (quadratic) difference tone to be added to the harmonics already there [1]. This distortion theory carried a long way, and was held well into the 20-th century by famous psychoacousticians like Georg von Békésy and Harvey Fletcher.

A totally new viewpoint on the phenomenon was offered by Jan Schouten, who based his Residue Theory on the then almost forgotten 19-th century observations of Seebeck. In his theory, the cause of the missing fundamental pitch phenomenon is the limited frequency resolution power of the human cochlea, specifically the basilar membrane. If a complex tone containing many harmonics is heard, the low-order harmonics, from fundamental up to about the 10-th harmonic, are essentially resolved by the cochlea and appear to the central nervous system as individual, separated pure tones. High-order harmonics, however, are only partially resolved or unresolved and appear to the nervous system as a cluster. This cluster contains information about the periodicity (or, equivalently, the fundamental frequency) of the sound in the form of its temporal envelope, which is processed by the central nervous system and causes the pitch sensation [2].

When I started research work in this area some 30 years later, Schouten's students and coworkers had already published quite a bit of experimental evidence on the missing fundamental pitch phenomenon in the literature. Some of the evidence, however, did not exactly support the residue theory behind it. For instance, RITSMA had found that the dominant region, that is, the range of harmonics that are most effective in conveying a fundamental pitch sensation, covers the 3-rd, 4-th and 5-th harmonics [3]. These harmonics are known to be resolved in the cochlea, and therefore should not cause a residue pitch sensation as defined by Schouten.

We started by proposing a new operational definition of pitch that reflected observed musical behavior closer than the conventional ANSI standard which defines pitch as "...that attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from low to high" [4]. The ANSI definition does indeed capture the ordinal properties of pitch sensations, but fails to capture its apparent ratio properties that are well known to musicians. Trained musicians not only can tell whether one note is higher or lower than another, they also can identify and name the musical interval that is defined by two successive or simultaneous notes. Even untrained listeners can usually tell when a performer sings out of tune, even if the general contour of the melody (i.e., the ordinal pattern of the notes) is correct. We therefore redefined pitch as an auditory sensation attribute that allows us to recognize a melody in a sequence of sounds, and designed our basic experimental paradigm, shown in Fig. 1, around the absolute identification of melodic (two successive note) intervals by properly trained musicians [5].
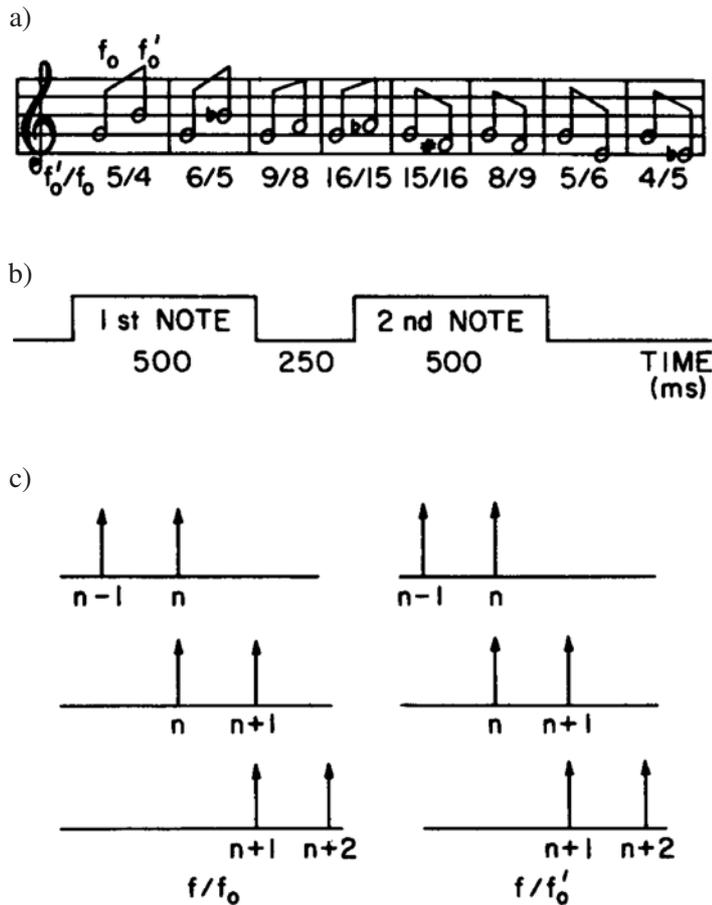
Fig. 1. Paradigm for musical interval identification experiments: (a) melodic intervals to be identified, (b) time-envelope of melodic-interval stimulus, (c) the three possible two-tone representations for each note. A random choice among the three possibilities was made for each note.

The results of our experiments, a sample of which is shown in Fig. 2, unveiled a couple of interesting things. The first finding was that, if two-tone complexes of successive harmonics were used to represent notes and correct interval identification score was viewed as a function of harmonic order, performance was always best for the lowest harmonic order and systematically deteriorated with increasing order. That behavior was not consistent with the residue theory, which predicted just the opposite. Furthermore, we found that musical interval identification behavior for monotic or diotic stimulus conditions (where both stimulus harmonics either go to one ear or to both ears) was essentially the same as for dichotic conditions (where each ear receives one stimulus harmonic). The dichotic condition, of course, makes the formation of a residue-based pitch impossible, since each cochlea receives only a single pure tone. These findings were interpreted as strong evidence that the residue concept was fundamentally wrong,

and that a missing fundamental pitch sensation must be the result of some central, neural process [5]. Consequently, most of the pitch models and theories that were developed afterwards [6–9] are central, neural theories.
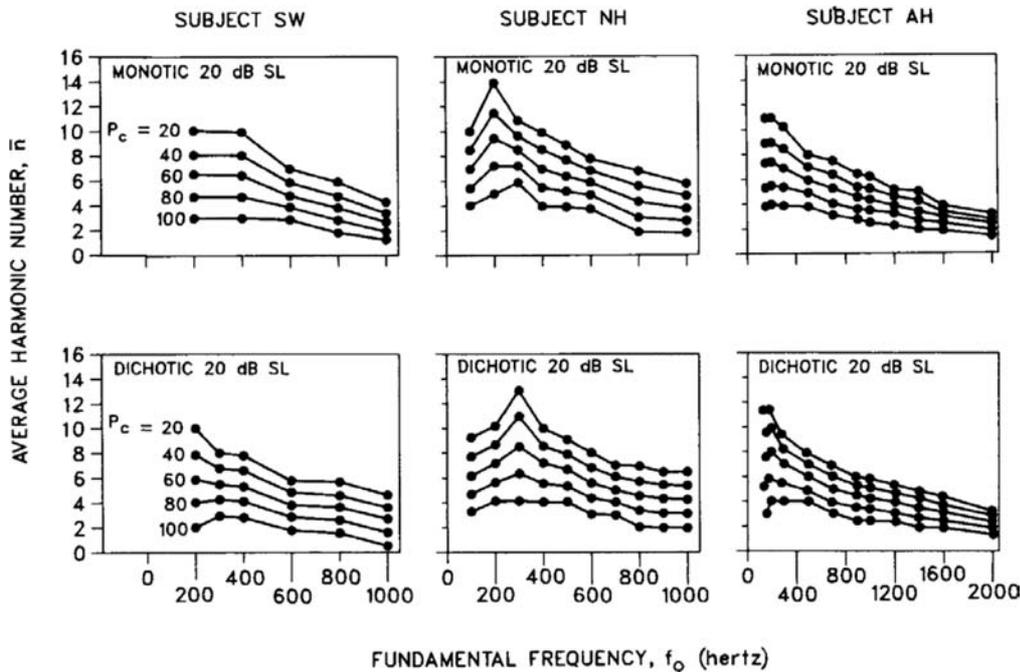


Fig. 2. Performance contours of three musically trained subjects, showing percentage correct identification, as a function of fundamental frequency of the first note (abscissa) and average lower harmonic number (ordinate). Top: monotic; bottom: dichotic.

As things often go when a pendulum swings, it may swing a bit too far as I soon discovered. After leaving M.I.T. and joining the Institute for Perception Research in 1982, I became involved in the study of prosody and pitch perception for speech sounds. In the course of that work I began to realize that a vowel, sung or spoken by a male voice, retains its pitch quality even if many low-order harmonics have been removed by highpass filtering. The highpass limit where pitch becomes difficult to identify is typically well beyond the 10-th harmonic. An important condition for pitch appeared to be that there are many high-order unresolved harmonics, not just two or three. In the late 1980s we were so fortunate to have Jacek Smurzyński, a former student of Andrzej Rakowski, in our laboratory for a year of postdoctoral research. Together we undertook a study that investigated our ability to identify and discriminate pitches of complex tones that contain many successive high-order harmonics.

We used stimuli comprising 11 successive harmonics, presented diotically, where the lowest harmonic could vary in order from 7 to 19. Using a musical (melodic) interval identification paradigm similar to the one shown in Fig. 1, we found that the score

was highest (100% correct) when the order of the lowest harmonic was 7, degraded to 60% correct when it was 13, and remained constant at 60% correct when the order increased further form 13 to 19. These results are illustrated in Fig. 3a. Using a simple 2-note (up-down) pitch discrimination paradigm, we found in a similar manner that the just-noticeable pitch difference (JND) increased from 0.8 Hz to about 5 Hz when the order of the lowest harmonic was increased from 7 to 13, and remained rather constant (between 5 and 6 Hz) when it was further increased from 13 to 19. Furthermore, the pitch discrimination experiment showed that JNDs for stimuli containing resolved harmonics are insensitive to changes in phase relations between harmonics, whereas JNDs for tones having only high-order unresolved harmonics can change by as much as a factor two when phase relations are changed [10]. This is illustrated in Fig. 3b. The conclusion from this work was that, as far as the missing fundamental pitch phenomenon is concerned, there appears to be a secondary pitch mechanism that operates on unresolved high-order harmonics when there are no low-order harmonics available. The pitch sensation it yields is much less salient than the pitch yielded by low-order resolved harmonics, and it needs many harmonics to have any effect. As far as modeling is concerned, there may be two separate and independent neural pitch mechanisms, one for resolved and the other for unresolved harmonics, or the phenomena we observed may simply be the results of one single neural mechanism that shows different behavior for aurally resolved and unresolved stimuli. This issue has, up to this day, not been decided.
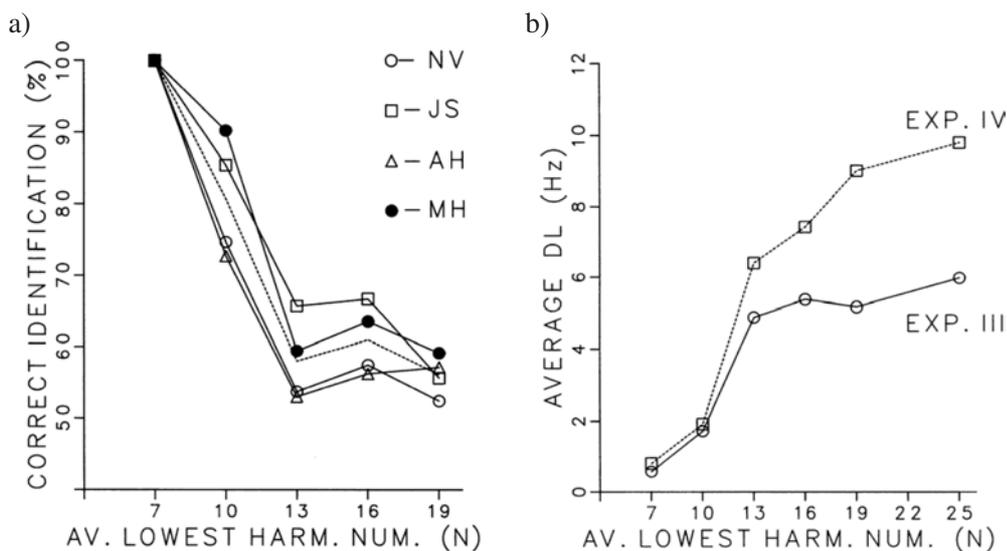


Fig. 3. Top panel: Melodic interval identification performance for four musically trained subjects, and their average score, plotted against the average lowest harmonic number of an 11-tone complex comprising successive harmonics of equal intensity. Bottom panel: Subject-averaged difference limens as a function of the average lowest harmonic number of the same 11-tone complex. For the solid function (Exp. III) tone components are added in sine phase, for the dashed function (Exp. IV) they are added in negative Schröder phase.

### 3. Harmonic intervals and chords

In another study we looked at the perceptual response to simultaneous complex tones, as happens in music when we listen to simultaneous voices (polyphony) or musical chords (homophony). Each note element of a polyphonic voice or musical chord can acoustically be represented by a complex tone that may or may not have energy at its fundamental frequency. The intriguing question then is to what extent, and how, our auditory system is able to separate all the perceived harmonics into the proper groups that correspond to each complex-tone element of a chord.

John Beerends did, as part of his dissertation, an experiment with musically experienced subjects who were given a response box with five buttons, representing the notes "do, re, mi, fa, so" (corresponding to the fundamental frequencies 200, 225, 250, 267 and 300 Hz). On each trial they heard two simultaneous notes chosen from this set, which could not include the unison (two identical fundamentals). Each of the two notes was independently represented by a pair of successive upper harmonics ranging from the $2^{nd}$ to the $10^{th}$ and chosen randomly. Subjects were asked to identify the perceived harmonic interval by playing it back on the button box, that is, pressing the perceived notes in sequential order. Four conditions were investigated, as illustrated in Fig. 4. In the first condition, thought to be the easiest, one two-tone complex (note) was presented to one ear, and the other two-tone complex (note) to the other ear. In this condition, the brain is given an extra spatial clue on how to partition the four perceived harmonics into the proper groups. In the second condition, all harmonics were presented to both ears, leaving it up to the brain to figure out how to partition the group. In the third and fourth conditions it was tried to mislead the brain by incorrect dichotic partitioning [11].
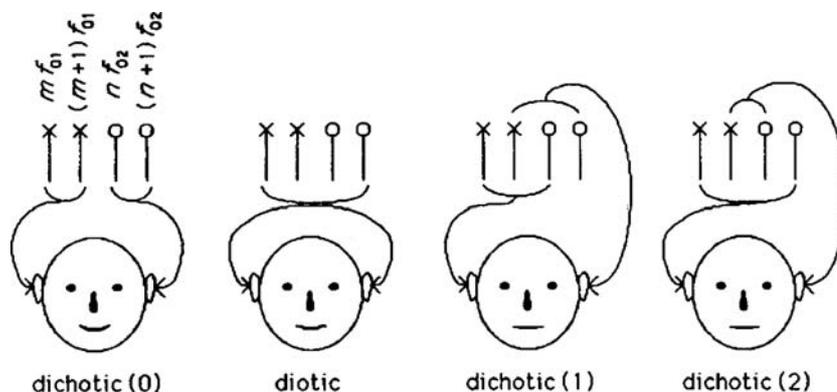


Fig. 4. Four different ways in which partials of two simultaneous two-tone complexes were distributed between subjects' ears.

The experiment yielded the somewhat surprising result that performance scores did not differ much between the four presentation conditions. Apparently our brain does not pay much attention in this task to the dichotic partition of stimulus partials, but seems to partition all available partials on the basis of harmonic relationships between perceived

frequencies. A second, less surprising result was that identification performance decreased rather monotonically from near-perfect when the harmonic order of both notes is 2, to chance level when it approached 10. A third, again surprising finding was that the probability of one of the two notes being identified correctly did not only depend on the harmonic order of its own representation, but also on the harmonic order of the other note. If the other note was of lower harmonic order (and therefore had a more salient pitch), the chance of being correctly identified was smaller than if the other note was of higher harmonic order. It seems that, to some extent, a note with a very salient pitch tends to inhibit the perception of the pitch of another simultaneous note that has a less salient pitch [11].

### 4. Pitch of carillon bells

Perfectly harmonic complex tones may be useful stimuli for the laboratory, but in musical practice they hardly occur. Musical instruments that, in theory, produce perfectly harmonic sounds (e.g., bowed violin, all reed or flue-type wind instruments), always show small deviations from perfect harmonicity that are caused by involuntary or deliberate actions of the player. Other melodic instruments, for instance the ones based on freely vibrating strings, show considerable amounts of inharmonicity. This is because actual strings, unlike the "ideal" textbook string, always exhibit some bending stiffness and their tension is not exactly constant during the oscillation cycle. Still other musical instruments, often found in the percussion section of a symphony orchestra, are designed to produce very inharmonic sounds, to the point where a pitch becomes difficult to recognize.

Inharmonic sounds are interesting to study from a pitch perception viewpoint. Where current pitch models readily agree on their predictions of the dominant pitch for a perfectly harmonic sound, they will typically differ when predicting the pitch of inharmonic sounds. Sounds of church bells and tuned carillons are popular study objects in the psychoacoustics and musical acoustics literature, because they typically contain a small group of partials with a (near) harmonic relationship, despite the instrument's inherent general inharmonicity. The Institute for Perception Research was only a few miles away from two major bell foundries (two of a total of four in the world that are capable of making tuned carillons), so it should not be surprising that the pitch of bells was part of its research program.

A bell is a freely vibrating three-dimensional structure that, when mechanically excited, will oscillated at natural frequencies that are determined by size, shape, thickness profile, and material. Since many degrees of freedom are involved, bells can, in principle, be made to produce almost any combination of pure-tone frequencies. Bells made and tuned according to European tradition typically produce an inharmonic tone series comprising the following main components with indicated relative frequencies: Hum (1), Prime (2), Third (2.4), Fifth (3), Octave (4), Twelfth (6) and Double Octave (8). The Third (a minor third), Octave, Twelfth and Double Octave are typically

strong components, whereas the Hum, Prime and Fifth are considerably weaker. The perceived pitch of such a bell corresponds with the frequency of the Prime, but does not seem to be controlled by this frequency. Both Hum and Prime can be mistuned considerable without really affecting the bell's perceived pitch. Until rather recently, it was not really known which bell partials mediate the pitch sensation, and what relationships exist between those partials and the perceived pitch.

LORD RAYLEIGH thought that the Octave partial was the only determining component, and that the perceived pitch was always one octave below this partial [12]. SCHOUTEN and 'T HART noticed that the Octave, Twelfth and Double Octave have frequency ratios of 2 : 3 : 4, and proposed the Residue Theory as explanation for the perceived pitch [13]. GREENHOUGH showed in a very brilliant but, unfortunately, hardly-known experiment with synthesized 8-partial bell sounds, that it is indeed the combined effect of the Octave, Twelfth and Double Octave partials that determines the perceived pitch [14]. In the late 1980s we performed an experiment similar to Greenhough's with real bell sounds, and obtained essentially the same results [15].

The sound of the C5 bell from the Hemony carillon of the church in Gouda, The Netherlands, was digitally recorded and stored in PCM code. Using digital filtering, each partial was separated, shifted in frequency by a certain amount, and added back. In this manner, 99 bell sounds were derived from the original, with the frequency of one of the lower nine partials shifted by amounts between 10% downwards and 10% upwards. (A 10% shift is about one full tone.) In a subsequent experiment, four musically-experienced listeners were asked to listen to an alternating sequence of bell sound and a pure tone and, for each processed bell sound, adjust the frequency of the pure tone till it matched the pitch of the bell. Results shown in Fig. 5, demonstrated unambiguously that only frequency shifts in the Octave, Twelfth and Double Octave partials affected the perceived pitch, as seen by the slopes of the matching functions. For all other partials, the slopes are essentially zero. The conclusion from this work was that, in order to contribute to pitch sensation, a partial must (1) be relatively strong and (2) have a frequency that bears a simple harmonic relationship with the expected pitch frequency. For instance, the Third is a strong component, but its relative frequency of 2.4 has no simple harmonic relation with the Prime (2.0). On the other hand, the Prime has a (unity) harmonic relationship with the perceived pitch, but is too weak to make an effective contribution to it.

One of the practical implications of these findings was demonstrated just last year in the development of the major-third carillon bell, an instrument that had first been introduced about a decade earlier [16]. These carillon bells have a major-third partial (relative frequency of 2.5) instead of the traditional minor third, which has important consequences for their sound and for the complexity of harmonies such carillons can play. Despite all positive features, there always was an acceptance problem among experts because of their weak "strike note." A bell's strike note is the attack portion of the bell sound that defines its pitch, as was illustrated in the right-hand panel of Fig. 5. It was thought for years that the decay time of partials was the culprit. Only recently it
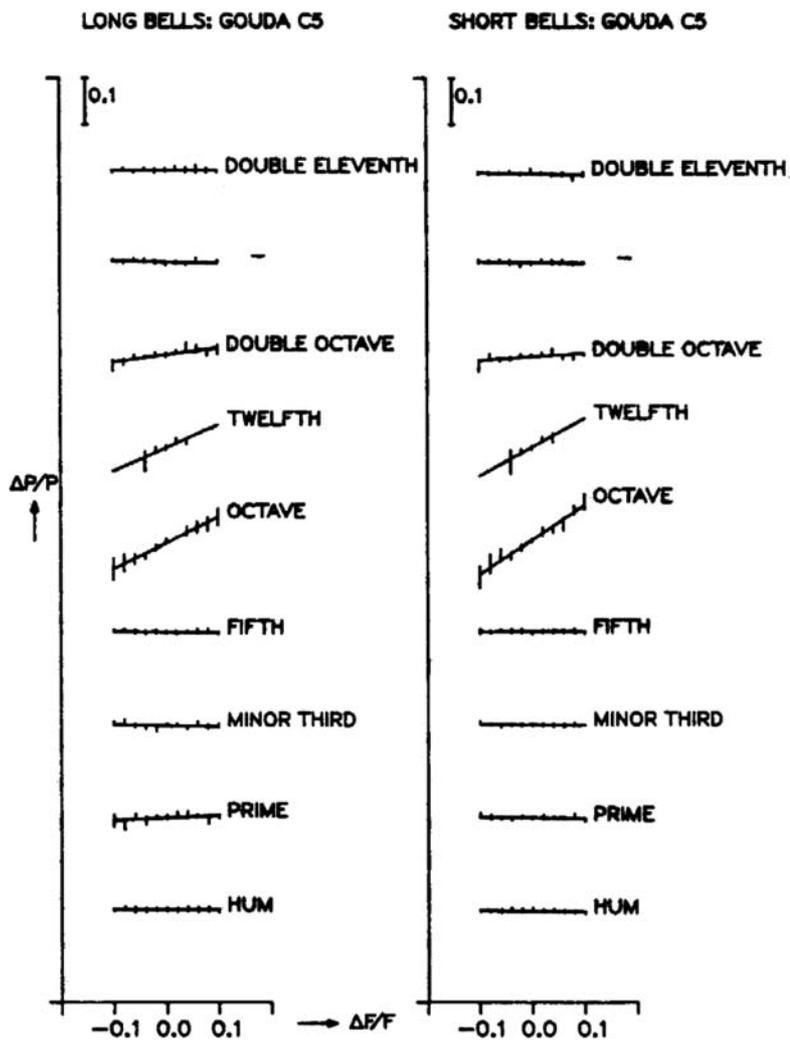
Fig. 5. Fractional shift of a bell's perceived pitch as a function of the fractional frequency shift of a specific partial. The left panel is for the first 1.5-s segment of sounds, the right panel when only the first 100-ms segment (strike note) is played. Error bars indicate the standard error of the means.

was found that the cause of the weak strike note was the tuning of the Twelfth, which was systematically too low by almost a quarter-tone. Fixing this tuning error solved the problem [17].

## 5. Prosody, the melody of speech sounds

At the Institute for Perception Research I worked in a research group on Hearing and Speech. This group had a long-standing interest and tradition in pitch perception,

which is no surprise since the group was originally formed by Jan Schouten. They had carefully followed the pitch work done at M.I.T. and had, in fact, developed a working pitch tracking model based on the Optimum Processor Theory of Goldstein [18, 7]. This model enabled them to extract and display the pitch of speech sounds as a function of time, known as the prosodic contour. It was a powerful research tool in the analysis and synthesis of speech, since it allowed researchers to make systematic changes in pitch contours and study their perceptual effects.

The details of the analysis technique and implications for intonation theory have been explained by 'T HART, COLLIER and COHEN [19]. The technique is based on a classical source-filter model of speech [20] which considers voice source (larynx) and voice filter (vocal tract) to be separable and independent components. The pitch tracker determines the voice source parameters for each running speech segment by deciding whether it produces noise or a periodic sound and, if found periodic, its rate. An LPC (linear predictive code) analysis tracks and maps the center frequencies and bandwidths of a number of formants, typically the first five. The power of such an analytic description is that measured pitch contours can now be systematically manipulated so that, after resynthesis, the sentence has a different prosodic contour but the voice character and the sentence's temporal and phoneme structure remain unchanged.

Studies of speech in various languages and dialects have revealed that there are very specific rules for intonation and prosody, but that these rules differ systematically from one language (and even dialect) to another. Using the above analysis-synthesis technique in combination with listening experiments by native speakers, intonation grammars have been developed for Dutch [21], German [22], British English [23], Russian [24], and even for English spoken by Dutch native speakers showing that learned intonation patterns tend to be transposed to other secondary languages [25]. Intonation was also found to be a dominant feature in the conveyance of emotions in speech [26].

The laboratory evidence that speech intonation and melody is an important perceptual feature for most languages should not come as a surprise. In some languages, known as "tonal" languages, the meaning of a word can change totally by changing its intonation. Infants learn to recognize intonation patterns in the parents' voice long before they learn the meaning of words. Playwrights and composers have exploited the power of human voice intonation for expression of emotion and drama in theater and opera for centuries. In the light of this it appears odd that intonation still receives so little attention in our methods of teaching foreign languages.

## 6. Beyond pitch perception

Pitch is without doubt an important sound attribute in music, but it is certainly not the only one. That is probably the reason why pitch studies led me to the investigation of other sound features that play a role in the production and perception of music or went well beyond the scope of music.

### 6.1. Perceptual entropy coding

In the early 1980s, Philips and Sony introduced the Compact Audio Disc (CD) which rapidly replaced the old vinyl records and set a new standard for audio quality. It soon became clear, however, that for music that is not played directly from a pre-recorded medium (e.g., radio, recordable disc, digital tape, internet sound), the CD bit rate (44100 samples/s × 16 bits/sample × 2 channels = 1.4 Mbits/s) could often not be met. In a successful attempt to retain subjective CD audio quality at bit rates much lower than the CD's, digital sound encoding schemes were developed that exploit our hearing system's masking features to hide unwanted coding artifacts (known as quantization noise) under the music signal so that they are not heard. When the first versions of these codes were announced, music critics and audio journalists were generally very skeptical and convinced that, with their well-trained golden ears, they would be able to hear the difference between an original CD sound and a bit-reduced version thereof. At the Institute for Perception Research, because of its Philips affiliation, it was considered a challenge to design a listening demo or test that would decisively convince such journalists of the contrary. We chose a blind adaptive two-alternative forced choice paradigm, using 2−3 s music samples taken from a CD, and sequentially presenting the original and a bit-reduced version in random order. Initially, the bit-rate reduction would be very large, for example, an average bit rate of 1 bit/sample (compared with 16 b/sample for the original), so that the difference would easily be heard and listeners would give a correct response. After two successive correct responses, however, the bit rate of the degraded signal was increased, making the difference harder to hear and increasing the probability of an incorrect answer. As soon as an incorrect response was given, the bit rate of the degraded signal was decreased, making the task easier again. With this "two-down, one-up" adaptation scheme, the trace stabilizes at a performance level of 71% correct, a reasonable indicator for a just-noticeable difference. This performance level was reached for most "golden ear" listeners when the degraded signal had an average bit rate of 2 bits/sample which, for a stereo signal, is equivalent to 176 kbits/s. The actual coding scheme that we used was later standardized as ISO-MPEG1-layer 1, a predecessor of MPEG1-layer 3 which is now popularly known as MP3. Other, even more sophisticated audio codecs have been developed since that time, allowing transparent coding (i.e., yielding results indistinguishable from CD sound) at rates down to 64 kbits/s [27]. As an alternative to human listening tests for the assessment of audio quality, the International Telecommunication Union has standardized a procedure that yields reliable predictions of perceived audio quality from simple physical sound measurements [28].

### 6.2. Noise-immune stethoscope

The Institute for Perception Research ceased to exist in the fall of 2001. A few months later I retired from the Technical University Eindhoven and went to work for the U.S. Army Aeromedical Laboratory at Fort Rucker, the home base of Army Aviation.

The helicopter environment is extremely noisy, and I had no idea of the multitude of problems for hearing preservation, communication, general health, and even survival that are caused by this noise.

Research projects done in this laboratory typically do not have an academic origin, but are inspired by very specific problems that Soldiers encounter during operations. Solutions must be found quickly because people's lives often depend on it. The temptation to resort to quick fixes is always there, and the real challenge is to use all available engineering, psychology and human factors skills to get to the heart of a problem and find principle-based, lasting solutions.

One of the problems for which we recently may have found a solution is auscultation of patients inside a helicopter. Helicopters are the backbone for the medical evacuation of wounded Soldiers from the battlefield to a field hospital. In the case of a chest wound which may have caused a pneumothorax or hemothorax, or if an endotracheal tube must be placed to help the patient breathe, a stethoscope is an almost indispensable instrument to help the treating physician make the right diagnosis and provide the proper treatment. Inside the helicopter, however, where such treatment is likely to be most critical, heart and breath sounds through a conventional stethoscope cannot be heard because the noise level can be as high as 120 dB(C). Even a modern electronic stethoscope in combination with insert earphones and helmet-mounted earmuffs does not work because most of the noise invades the stethoscope through its mechanical/acoustical sensor. Even turning up the volume does not help, since this only increases the stethoscope's output level but does not improve its signal-to-noise (S/N) ratio.

The development of a new type of stethoscope that could deal with the extreme noise levels inside a helicopter took place in several stages. The basic design of a modern electro-mechanical stethoscope, where tissue-borne vibrations are mechanically transmitted to a piezoelectric transducer and transformed into an electrical voltage signal, served as a starting point. First, an attempt was made to construct the mechanical part of the sensor in such a way that its mechanical input impedance would be close to that of the human chest wall. If impedances are close, only a small fraction of the energy from a tissue-borne vibration signal will reflect at the flesh-sensor boundary, and most of it will be transmitted into the sensor. Air-borne noise, on the other hand, will be mostly reflected because the equivalent mechanical impedance of air is much smaller than that of the chest wall or the sensor. This technique yielded an instrument that, in a noisy environment, performed better than most other electronic stethoscopes, but still fell far short when it came to beating the noise level of a flying helicopter [29].

A next attempt to improve S/N ratio of the device was to add noise barrier rings around the sensor, to stop and reflect waves that are excited by the noise field in the tissue of the patient and propagate along the skin as surface waves [30]. This feature increased the S/N ratio by another 5 dB, but still not enough to make the device effective in a helicopter environment. It seemed that we had stretched the conventional electro-mechanical transduction technology about as far as it would go, and that a totally different set of physical principles and associated technology would have to be included to make the instrument work.

That technology turned out to be ultrasound, a branch of acoustics that muscial and psychoacousticians usually don't deal with. It works as follows. A high-frequency (2-3 MHz) sinusoidal sound carrier wave is emitted into the patient's chest from an ultrasound transmitter, which typically consists of a wafer made of piezoelectric material (like lead- zirconate or barium-strontium titanate) and is made to oscillate by an electrical signal input. When this wave meets a tissue boundary, part of the carrier wave will be reflected because of the sudden impedance difference. Moreover, if this tissue boundary moves toward the stethoscope the reflected wave will have a slightly higher frequency than the carrier, caused by the Doppler effect. Conversely, if it moves away from the transmitter, the reflected frequency will be lower. Therefore, the velocity of the moving tissue boundary is frequency-modulated on the reflected carrier, and can be made audible by simple demodulation, just as is done in a FM radio receiver. The reason why this technique is essentially noise immune is that the stethoscope communicates with the patient's body in a very high frequency band where the helicopter does not produce any noise. This technique works, of course, equally well inside an ambulance or in a crowded and very noisy sport stadium.

The new stethoscope, which is a hybrid combination of a conventional electromechanical and ultrasound-Doppler operation, has been tested in background noise levels up to 120 dB SPL and was found to produce clean, noise-free images of heart and lung sounds [31, 32]. Figure 6 shows an example of a heartbeat image recorded in broadband noise with a Littmann Cardiology III conventional stethoscope coupled to a Brüel & Kjær microphone (a) and with the hybrid stethoscope operating in its ultrasound mode (b), at noise levels of 100 and 110 dB(C), respectively.
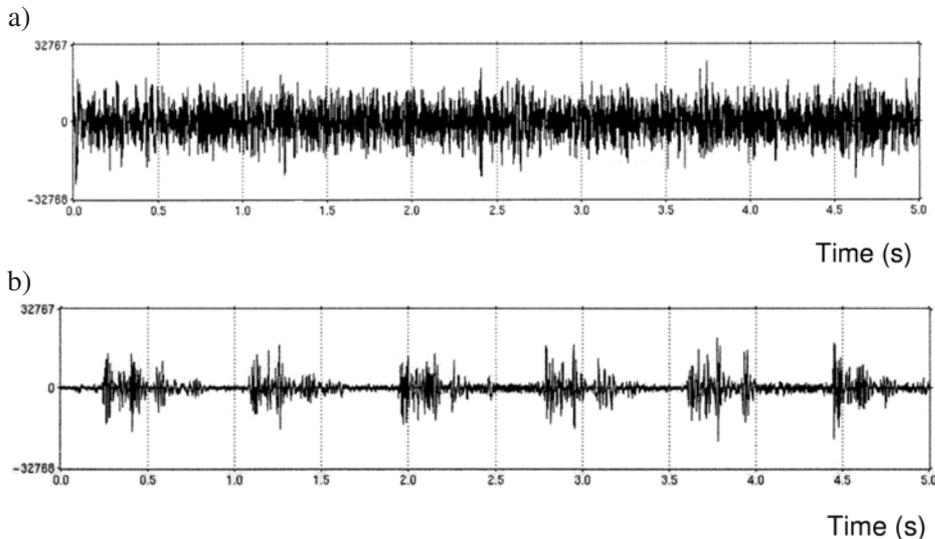


Fig. 6. The audio signal from a human heart beat with 100-dB(C) background noise observed through a Lttmann Cardiology III acoustical stethoscope (a), and with 110 dB background noise observed through an ultrasound-Doppler stethoscope (b).

The device is currently undergoing clinical testing to investigate whether the noise-free sound produced by ultrasound imaging (see Fig. 6b) has the same diagnostic potential as sound produced by conventional auscultation methods. Because the physical principles underlying the technologies are so different, it should be expected that a heartbeat, as heard through a conventional stethoscope, sounds different when it is heard through an ultrasound-Doppler device. This is indeed found to be the case, and will probably necessitate additional training for physicians and medics to learn the relations between new sounds and specific heart and lung anomalies or injuries.

## 7. Epilogue

I vividly remember my first meeting with Andrzej Rakowski. It was in the early 1970s in Cambridge, Massachusetts, in my office at MIT. I had just completed my PhD program and was struggling as a young staff member in the departments of Music and Electrical Engineering to develop new courses and to publish my first papers. Andrzej came for an orientation visit while on tour in the USA, and had been so kind to include me in his schedule.

I remember talking mostly about my thesis work, which dealt with the phenomenon of the missing fundamental and our finding that a musical pitch sensation can be evoked by two dichotically presented successive harmonics. Andrzej, in turn, showed me some of the pitch perception work that he had done in his Warsaw laboratory, particularly the very small values for just-noticeable difference in tone frequency that he had been able to obtain with highly-trained listeners. What especially intrigued me about this work was its setting, the Fryderyk Chopin Academy of Music, which is the main music conservatory in Warsaw. It was the beginning of an eye-opener for me to learn that research does not only serve an academic purpose, yielding general knowledge and quantitative models, but that hearing research can also have an important impact on musical practice such as composition, performance, education, development of new musical instruments, and a host of other things as I would later discover.

Since that first meeting, it took more than 20 years until the political climate between East and West had improved enough to allow regular visits, open exchanges of ideas and actual cooperation on projects. In 1986 I made my first trip to Poland, which included a visit to Andrzej's lab. More visits took place in the following years. I had moved to the Institute for Perception Research at the Technical University of Eindhoven in the Netherlands, from where it was easier to set up and maintain a steady working relationship. Especially after the Wall came down in 1989, we were able to invite colleagues in Eastern Europe for work visits and postdoctoral fellowships, and Andrzej's laboratory became part of this exchange.

Looking back at all the projects and activities in the field of acoustics in which I have been involved, I am glad to have spent so much time and effort in the beginning on trying to understand fundamental issues and principles. Later on this has paid off many times when I encountered seemingly new problems and realized that there often is an old solution. An example is the surface wave problem that we encountered with electro-mechanical stethoscopes. The process causing this problem is the reverse of

what is known from vibro-tactile sensory perception experiments, where the technique of preventing a stimulus from propagating away along the skin by using a "surround" has been known for quite some time. Andrzej Rakowski has been an inspiring example and role model for this fundamental attitude in science which he was able to transfer to many of his students and colleagues. I saw for myself to what innovative applications this had led him when he took me to the Warsaw Opera House on my first visit to Poland. He had designed an acoustically very effective stage setting for the garden scene in La Traviata using optically transparent but acoustically reflective plexiglass. Needless to say, the performance was unforgettable.

## Disclaimer

Opinions, interpretations and conclusions contained in this article are those of the author, and are not necessarily endorsed by the U.S, Army and/or the Department of Defense.

## References

[1] HELMHOLTZ H. L. F. VON., *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*, F. Vieweg & Sohn, Braunschweig 1863.

[2] SCHOUTEN J. F., *The residue and the mechanism of hearing*, Proceedings of the Koninklijke Nederlandse Academie voor Wetenschap **43,** 991–999 (1940).

[3] RITSMA R. J., *Frequencies dominant in the perception of the pitch of complex sounds*, Journal of the Acoustical Society of America, **42,** 191–198 (1967).

[4] American National Standards Institute, *American national psycho-acoustical terminology*, S3.20, New York 1973.

[5] HOUTSMA A. J. M., GOLDSTEIN J. L., *The central origin of the pitch of complex tones: Evidence from musical interval recognition*, Journal of the Acoustical Society of America, **51,** 520–529 (1972).

[6] TERHARDT E., *Zur Tonhöhewahrnehmung von Klängen II: Ein Funktionsschema*, Acustica, **26,** 187–199 (1972).

[7] GOLDSTEIN J. L., *An optimum processor theory for the central formation of the pitch of complex tones*, Journal of the Acoustical Society of America, **54,** 1496–1516 (1973).

[8] BILSEN F. A., GOLDSTEIN J. L., *Pitch of dichotically delayed noise and its possible spectral basis*, Journal of the Acoustical Society of America, **55,** 292–296 (1974).

[9] MEDDIS R., HEWITT M. J., *Virtual pitch and phase sensitivity of a computer model of the auditory periphery, I. Pitch identification*, Journal of the Acoustical Society of America, **89,** 2866–2882 (1991).

[10] HOUTSMA A. J. M., SMURZYNSKI J., *Pitch identification and discrimination for complex tones with many harmonics*, Journal of the Acoustical Society of America, **87,** 304–310, 1990.

[11] BEERENDS J. G., HOUTSMA A. J. M., *Pitch identification of simultaneous diotic and dichotic two-tone complexes*, Journal of the Acoustical Society of America, **85,** 813–819 (1989).

[12] RAYLEIGH LORD, *On bells*, Philosophical Magazine, **29,** 1–17 (1890), Reprinted [in:] *Acoustics of Bells*, ROSSING T. D., [Ed.], Van Nostrand Reinhold, New York 1984.

[13] SCHOUTEN J. F., 'T HART J., *The strike note of bells*, Journal of the Acoustical Society of the Netherlands, **7,** 8–19 (1965).

[14] GREENHOUGH M., *The analysis and synthesis of bell sounds*, Proceedings of the Institute of Acoustics, September (1976).

[15] HOUTSMA A. J. M., EGGEN J. H., *Indian drums and Chinese bells: Measured spectrum and perceived pitch*, [in:] E. OZIMEK [Ed.], Subjective and objective evaluation of sound, World Scientific Press, Singapore 1990.

[16] SCHOOFS A. J. G., ASPEREN F. VAN., MAAS P., HOUTSMA A. J. M., THOLEN H. J. G. M., LEHR A., *A carillon of major-third bells (I, II, III)*, Music Perception **4,** 3, (1987).

[17] LEHR A., *The significance of the twelfth for the formation of the strike note* [in Dutch], www.andrelehr.nl, 2006.

[18] DUIFHUIS H., WILLEMS L. F., SLUYTER R. J., *Measurement of pitch in speech: an implementation of Goldstein's theory of pitch perception*, Journal of the Acoustical Society of America, **71,** 1568–1580 (1982).

[19] 'T HART J., COLLIER R., COHEN A., *A perceptual study of intonation: An experimental-phonetic approach to speech melody*, Cambridge University Press, Cambridge 1990.

[20] FANT C.G.M., *Acoustic theory of speech production*, Mouton, The Hague 1960.

[21] COLLIER R., 'T HART J., *A grammar of pitch movements in Dutch intonation*, IPO Annual Progress Report **6,** 17–21 (1971).

[22] ADRIAENS L. M. H., *Ein modell deutscher intonation*, Dissertation, Technical University Eindhoven, 1991.

[23] ELSENDOORN B. A. G., *Tolerances of durational properties in Britisch English vowels*, Dissertation, Technical University Eindhoven, 1984.

[24] ODÉ C., *Russian intonation: A perceptual description*, Dissertation, Leiden University, 1989.

[25] WILLEMS N. J., *English intonation from a Dutch point of view*, Dissertation, Technical University Eindhoven, 1982.

[26] MOZZICONACCI S., *Speech variability and emotion*, Dissertation, Technical University Eindhoven, 1998.

[27] HOUTSMA A. J. M., *Perceptually-based audio coding*, [in:] VORLÄNDER M., KUWANO S., HAVELOCK D. [Eds.], Handbook of signal processing in acoustics, Springer, New York 2007 (in press).

[28] International Telecommunications Union, *Method for objective measurement of perceived audio quality*, ITU-R.BS 1387, 2001.

[29] CAIN P. A., AHROON W. A., *Auscultation in noise: A program to develop a stethoscope capable of functioning during aeromedical transport*, Journal of the Acoustical Society of America, **112,** 2274 (2002).

[30] HOUTSMA A. J. M., *High noise environment stethoscope*, U.S. patent application nr. 11/425312, 2006.

[31] HOUTSMA A. J. M., CURRY I. P., SEWELL J. M., BERNHARD W. N., *A dual-mode noise-immune stethoscope for use in noisy vehicles*, Proceedings Army Science Conference 2006, paper K-27, Orlando, 2006.

[32] HOUTSMA A. J. M., CURRY I. P., SEWELL J. M., BERNHARD W. N. *Auscultation in high-noise environments using hybrid electromechanical and ultrasound-Doppler techniques*, Journal of the Acoustical Society of America, **120,** 3361 (2006).