BINAURAL ANALYSIS METHODS AND THEIR RELATIONSHIP TO QUALITY EVALUATION OF HANDS-FREE TELECOMMUNICATION EQUIPMENT

H. W. Gierlich
HEAD acoustics GmbH, Ebertstr. 30a 52134 Herzogenrath, Germany
Tel.: +49 2407 57722; Fax: +49 2407 57799

Abstract

Since modern telecommunication equipment, especially hands-free telephones, incorporates sophisticated signal processing, the analysis methods must take into account the properties of the human hearing. The basis for the correct aquisition of test data -used for auditory but instrumental measurements as well- is the binaural rcording and binaural analysis of the test stimuli. The paper gives an overview, in what ways binaural methods can be applied for quality evaluation. The paper focusses on methods for aquiring test data in the listening situation, in the converational situation and for instrumental measurements using defined, artificial test stimuli. Various methods for playback of binaurally recorded sounds in different situations are shown.

1 INTRODUCTION

Enhanced signal processing in modern tele-communication equipment allow the use of such terminals in various acoustical environments and under much worse signal to noise conditions than typically thought of in the past. The signal processing used is typically highly time variant and non stationary. The use of standard measurement technology known in telephonometry for determining the transmission characteristics is no longer possible. Since a complete instrumental evaluation of new equipment is still not possible completly, first the auditory evaluation of such equipment needs to be made using hearing tests and conversational tests in order to find instrumental evaluation procedures. Listening tests are used to determine the auditory perceived quality of the hands-free sets when being in the listening only situation. The artificial head is used for sound generation as well as for sound reception. This ensures the correct directivity of the sound generation and sound reception at the hands-free side and on the handset side in case a handset is used in combination with a hands-free setup.

2 THE PRINCIPLES OF BINAURAL TECHNOLOGY

2.1 Free Field Conditions

Human hearing analysis the input signals to the two ears with respect to both interaural time delay and directional filtering effected by the outer ear. In the horizontal plane, mainly the interaural time difference is important in detecting the direction of sound incidence, but also correct filtering of input signals to both ears is essential. In the median plane, correct filtering is especially significant due to the lack of interaural time differences and must recognize direction by outer ear filtering only, provided by the elements head and shoulder. The head transfer function changes considerably, depending upon the direction of sound incidence. Moreover, the head transfer function differs from person to person, so that an 'average head transfer function' is necessary for designing of artificial heads with average transmission characteristics. A definition of the required properties both, acoustically and geometrically can be found in [1,2].
The physical effects which determine the head transfer functions are diffractions, resonances and reflections as caused by the acoustically relevant elements: head, torso, shoulder, pinna, cavum conchae, ear canal and ear drum, etc.. These effects can be divided into direction-dependent and direction-independent ones. Since people possess these acoustically relevant elements, the idea of averaging the geometric data of different test subjects suggests itself. To achieve this, assessment of the correct position of the elements is obviously essential, because incorrect positioning affects the direction-dependent component of the head transfer function and cannot be corrected later.
A model describing the principal influence of the different acoustically relevant elements is described in [3]. The model differentiates between directional

dependant and directional independant components. Directional dependant diffraction effects due to the torso leads to increases or decreases of ± 3 dB below 1 kHz. The shoulder diffraction leads to increases or decreases of about ± 5 dB in the frequency range below 2 kHz. The distance between the shoulder and the ear-canal entrance determines the magnitude and location in frequency. The influence of the basic head structure is strongly directional as well. It In the horizontal plane, the signal is lowpass-filtered with a cut-off frequency of about 1 kHz. More detailed information can be found in [3]. Level variations of as much as - 30 to + 15 dB occur in the frequency range above 1 kHz as the sound source is moved from the far to the near side. Because the ear-canal entrance is not located exactly in the middle of the head, a comb filtering effect occurs, depending upon the direction of sound incidence. The influence of the pinna is both direction-dependent and direction-independent. The cavum conchae contributes a broadband amplification of 20 dB maximum in the frequency range of about 1 to 10 kHz, independent of direction. Diffraction effects at the pinna result in peaks and troughs in the head transfer function above about 2 kHz, depending upon direction.The influence of the ear canal is direction-independent if far-field application is assumed.

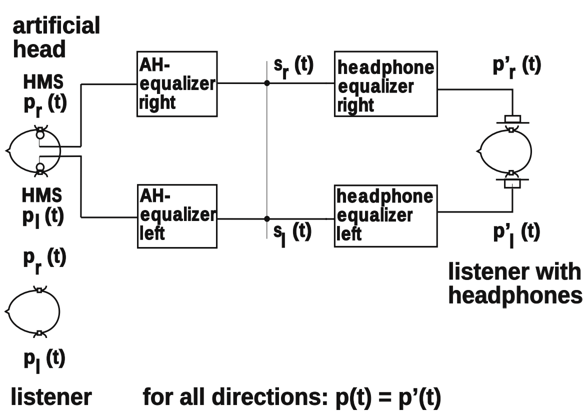

Fig.1: Principle of artificial head recording and reproduction

Since the transfer characteristics of artificial heads is strongly directional dependant, the output is not comparable to a standard microphone. An equalization is required in order to achieve for reference conditions (e.g. freefield, diffusefield) compatible output signals to a standard measurement microphone as well as to achieve loudspeaker compatibility for binaural recordings (see [3],[5]). Fig. 1 shows, how typically artificial heads are setup for recording and reproduction of sounds.

2.2 Close to Ear Sound Sources

For sound sorces close to the ear the situation is different. In this situation the acoustical impedance as well as the compliance of the human pinna need to be taken into account in order to achieve the transfer characterisics between handset and ear than for the average user. A simplified pinna simulation [3, 6, 7], is based on mathematically describable elements. The modification described in [6] allows the application of this pinna for freefield as well as for close to ear sound sources. Since the compliance of this simplified pinna is similar to typical human pinnas, this pinna is excellent applicable for handset measurements with realistic pressure forces. A complete description can be found in [6]. The typical pressure force dependant transfer characteristic of a handset, measured using the simplified pinna, is shown in Fig.2

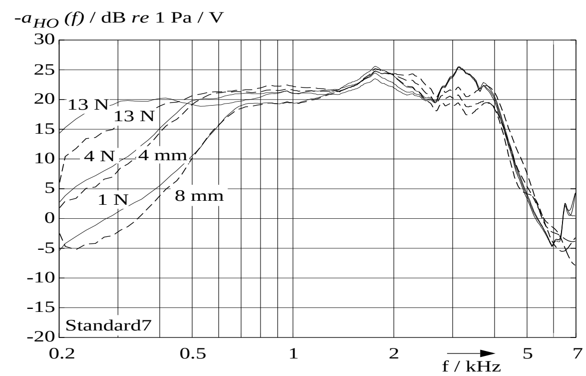

Fig.2: Typical [6] pressure force dependant transfer function for the receiving direction of a handset

3 RECORDING OF TEST STIMULI FOR AUDITORY AND INSTRUMENTAL TESTING

The most obvious application is the use of artificial heads tor the recording of test stimuli esp. for hearing tests [8, 9]. Since the use of equalized headphones guarantees the reproduction of sounds as close as possible to the original sound situation it is certainly a possibility to record signals for listening tests using artificial heads. However it has to be discussed in advance, what impact the use of headphone has for the experiment conducted. The advantage of being able to reproduce the same hearing event at any time comes with the disadvantage of „unnatural“ use of headphones. If this situation is explained to the test subjects adequately, the results gained in such listening tests are certainly as valid as such while the test persons use a real hands-free set for the tests. In any case tests are repeatable with a much higher accuracy since the virtual acoustical environment as well as the recorded samples are much more stable

than real time experiments, esp. in the case of non linear and/or time variant hands-free telephones.
A typical example of a recording setup is shown in Fig. 3.

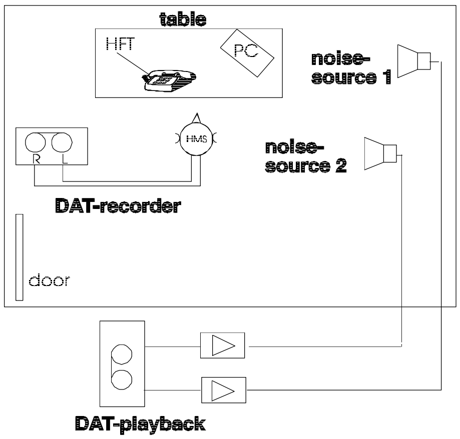

Fig.3: Recording setup for listening tests using an artificial head with hands-free telephones [8]

A sound situation including noise sources is recorded and stored on a DAT-tape. Since typically the noise sources are timevariant, such a situation can be (during the recording) well synchronized to the different source signals (speech) in order to ensure a constant masking during the whole test suite. When emitting sounds by the artificial mouth while listening the same time also double talk situation can be simulated for listening tests the same time. Of course the test persons need to be instructed adeqately, since during the tests they might be listening to their „own voice" without speaking. The advantage of such simulated double talk tests is, that the test setup is a listening test. In such a situation it is much easier to concentrate on the double talk while the test subjects are not speaking themselfs.
In general the same advantages exist if a handset is used instead of a second hands-free telephone. Such a test setup (as shown in Fig. 4) allows the evaluation of a very typical situation (hands-free set in combination with a handset telephone). In such a situation the listener is much more sensitive to quality problems of the hands-free telephone since no masiking by room noise or reverberance is present. In this situation very typical conditions can be recorded in a controlled test setup. The leakage effect of the handset can be included in the experiements by using realistic pressure forces applied to the handset while using a flexible pinna simulation as described in 2.2. A room noise injection is possible as well recording defined noise sources even at different locations. By definition the exact amount of room noise is transmitted by the acoustical leakage (between handset and ear), by electrical sidetone and on the ear not covered by the handset [6, 10].

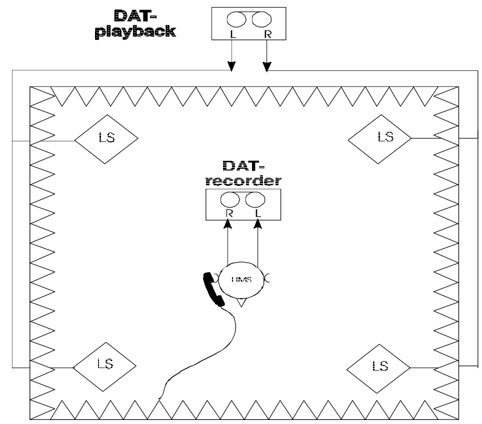

Fig.4: Recording setup for listening tests using an artificial head with handset and a 4-loudspeaker setup for room noise simulation using binaurally recorded noise sequences

A very interesting possibbility for room noise simulation is shown in Fig. 4. Four loudspeakers are used for playback of noise which has been prerecorded binaurally. Such a playback arrangement allows a virtual display of the recorded noise sources with similar accuracy than headphone reproduction [11].Thus a room-related plaback can be achieved which allows a room noise simulation under controlled conditions with a high accuracy.

4 EXAMPLES OF AUDITORY EVALUATIONS

For a situation where a hands-free telephone was connected to a handset telephone some examples of auditory evaluations are given. The tests were conducted for unidirectional transmission. During the listening tests various attributes were used to evaluate the „quality" of the hands-free set. Typical attributes, which determine the quality of a hands-free set in the listening only situation are: quality of room noise transmission, telephone speech quality, completness of speech transmission, sound, clipping, reverberation, noise. To evaluate these parameters, a testsetup similar to the one described in 2 was chosen. For the auditory judgement of the binaurally recorded situations both, hands-free and handset, headphone reproduction was used [9]. In Fig. 5 the mean opinion scores for four different hands-free telephones are given for the attribute „telephone speech quality" in the receiving direction of the hands-free telephones. Speech levels are chosen to -4,7 dBPa at the mouth reference point. Clear differences can be seen for the different sets. Set 2 reaches an MOS of 3.5, it sounds relatively loud and clear with good speech intelligibility. Set 1 reaches only MOS 2, it sounds distorted and unnatural. Set 3 and 4 both sound dull and bandlimited and even distorted. However the distortion is less audible than for set 1.

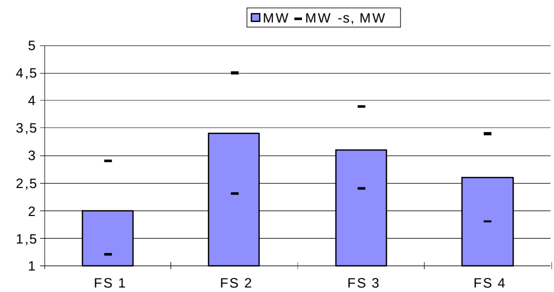

Fig.5: Listening opinion test, MOS by evaluating the „telephone speech quality“ in the receiving direction, ambient room noise level 40 dBA

The same test but for the sending direction was also carried out. The test signals were recorded binaurally using a handset positioned with a defined pressure force to the artificial ear. The results, shown in Fig. 6 make obvious that the auditory judgement is different for the sending direction of the hands-free telephones. Set 1 and 2 now achieve MOS around 3 due to their relatively balanced and clear sound. Set 3 and 4 sound dull and bandlimited, set 3 in addition produces audible distortion.

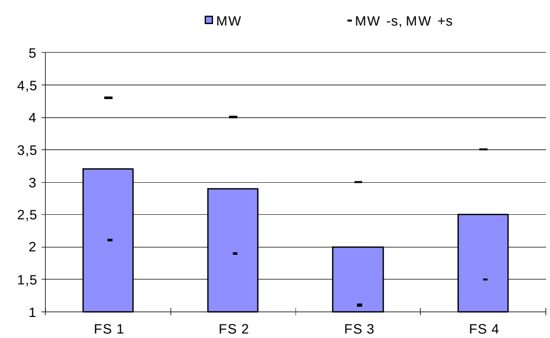

Fig.6: Listening opinion test, MOS by evaluating the „telephone speech quality“ in the sending direction, ambient room noise level 50 dBA

5 EXAMPLE OF INSTRUMENTAL EVALUATIONS

An example for instrumental analysis of hands-free telephones is shown in Fig. 7. For evaluation a Composite Source Signal [12] is used, which to a certain extent is speech-like. In the spectral representation (Fig. 7) time domain as well as frequency domain effect become obvious. Bandlimitation, a high- (500 Hz) and low-pass filter response (2,5 kHz) can be seen. In addition a time and frequency variable comb-filter structure becomes visible. No clipping effects can be seen. This analysis correlates well to the auditory measurements.

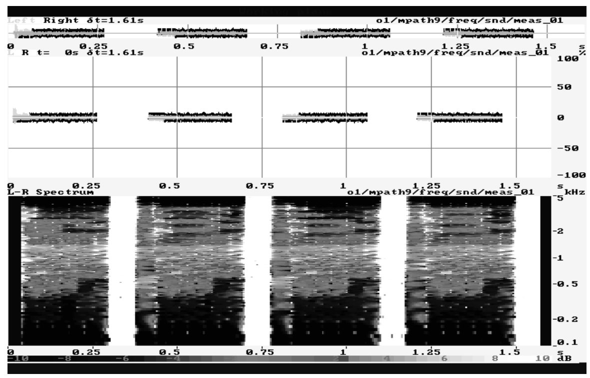

Fig.7: HFT 3, sending, CS-signal measurement
upper: time signal (dark: exitation signal, light: measured send signal), lower: FFT-spectrography (measured send signal/ FFT exitation signal)

Part of the work was supported by Deutsche Telekom.

References

[1] ITU-T, Recommendation P.58, 1994
[2] ANSI, S. 3.36-1985 (ASA S8-1985): Specifications for simulated in-situ airborne acoustic measurements, Standards Secretariat, Acoust. Soc. Am.
[3] Genuit, K.: Ein Modell zur Beschreibung von Außenohr-übertragungseigenschaften, Diss. RWTH Aachen, 1984
[4] Blauert, J.: Spatial Hearing, MIT Pres, Cambridge, 1983
[5] Gierlich, H.W., Genuit, K.,: Processing artificial-head recordings, J. Audio Eng. S., Vol. 37, No. 1/2,(1989), 34-39
[6] Krebber, W.: Sprachübertragungsqualität von Fernsprech-Handapparaten, Diss. RWTH Aachen 1995, VDI R 10,19
[7] Gierlich, H.W., Krebber, W.; Zur Reproduzierbarkeit von Messungen einer neuen vereinfachten Pinna- Simulation, DAGA 1995, pp. 667-670
[8] Gierlich, Euskirchen, Hottenbacher, Kettler: Analyse der Übertragungsqualität von Freisprecheinrichtungen; DAGA 95, pp.671-674
[9] Kettler, Hottenbacher, Heinrichs, Gierlich: Verfahren zur subjektiven Qualitätsbeurteilung von Freisprecheinrichtungen, DAGA 1994
[10] Gierlich H. W.: Sprachverständlichkeit beim Fernsprechen unter Störschalleinfluß mit besonderer Berücksichtigung von Freisprecheinrichtungen, Diss. RWTH Aachen 1988, Der Fernmeldeingenieur Heft 8-10
[11] Xiang, N. Genuit, K., Gierlich, H.W.: Investigations on a new reproduction porcedure for binaural recordings, AES New York, 1993, Preprint 3732
[12] Gierlich, H.W.: A measurement technique to determine the transfer characteristics of hands-free telephones, Signal processing, Vol.27, No.3, 1992, pp. 281-300