Practical Limitations of Wideband Terminals

Dr.-Ing. Carsten Sydow
Siemens AG
ICM CP RD VD1
Grillparzerstr. 12a
81675 Munich, Germany
E-Mail: sydow@siemens.com

Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction
8th and 9th June 2004 – Mainz Germany

Introduction

Wideband telephony was standardized for ISDN terminals in the 1990s by ETSI [2][3], but never found it's way into widespread deployment. The reasons were the price for such telephones and the fact, that always both parties have to have a wideband phone in order to establish a wideband call.
Nowadays PCs can be found in nearly every home and the hardware like soundcard and headset is well capable of transmitting wideband telephony signals. That's why for IP-telephony, wideband speech transmission comes back into the focus. Also non-PC IP telephone terminals start to support wideband speech transmission like the Siemens "optiPoint 410" IP-phone [9].
While a variety of high-quality codecs are available for compressing the wideband speech for transmission like the G.722. G.722.1, G.722.2 (AMR wideband) and iSAC [8] codecs, the transducers used in common telephone terminals are usually not suited to transmit or receive the whole bandwidth of the wideband speech signal. The article presents the requirements posed by the wideband standards and discusses the problems to meet the standards with conventional transducers.

Standards for Wideband Telephony

Applicable recommendations for digital narrowband telephones, ISDN or IP phones, are, among others:

operating mode	country	narrowband	wideband
handset	Europe	TBR8 [4]	I-ETS 300 245-5 [2]
	USA	TIA 810A [5]	TIA 920 [6]
hands free	Europe	I-ETS 300 245-3 [1]	I-ETS 300 245-6 [3]
	USA	TIA 810A [5]	TIA 920 [6]

Main differences between narrowband and wideband requirements are the extended frequency range, the different artificial ear used for handset measurements. Minor differences can be found in the distortion and noise requirements.

Frequency Masks

Figures 1 to 4 show the different frequency masks for handset and hands free operating modes in the send and receive direction.

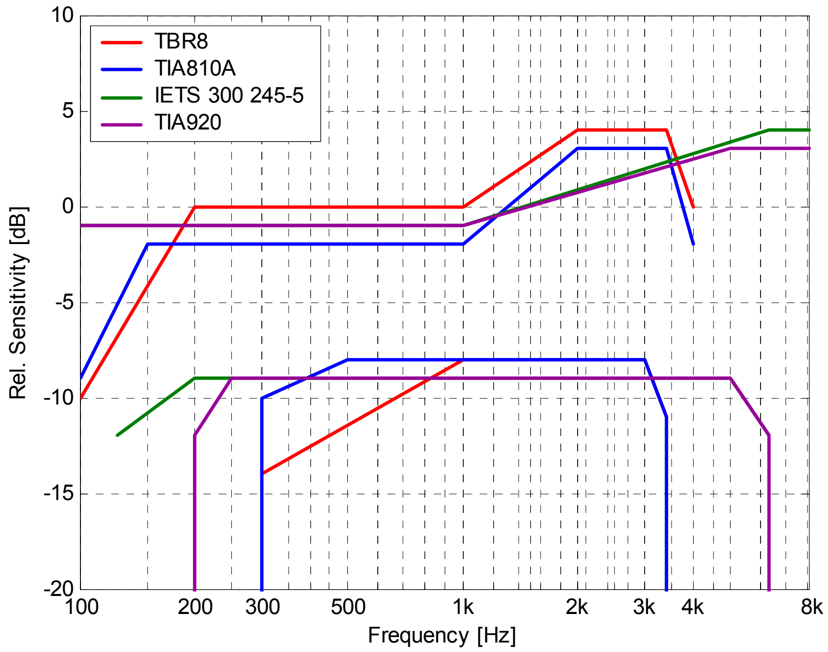

Figure 1: Handset sending frequency response masks

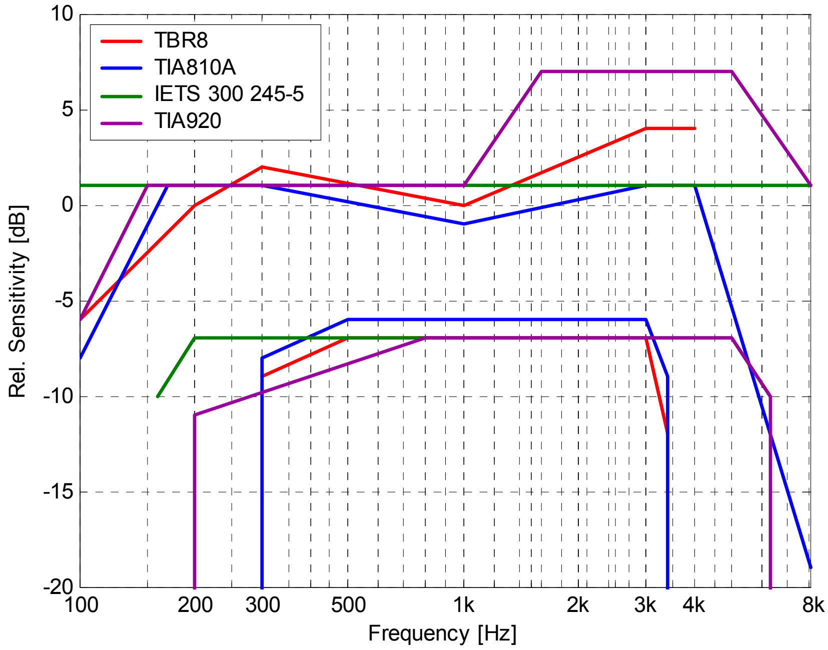

Figure 2: Handset receiving frequency response masks

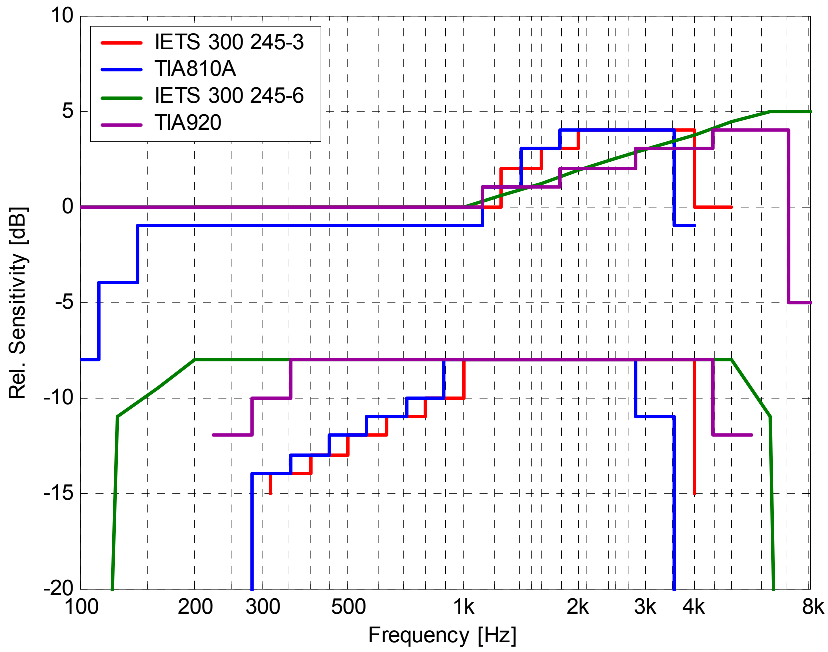

Figure 3: Hands free sending frequency response masks

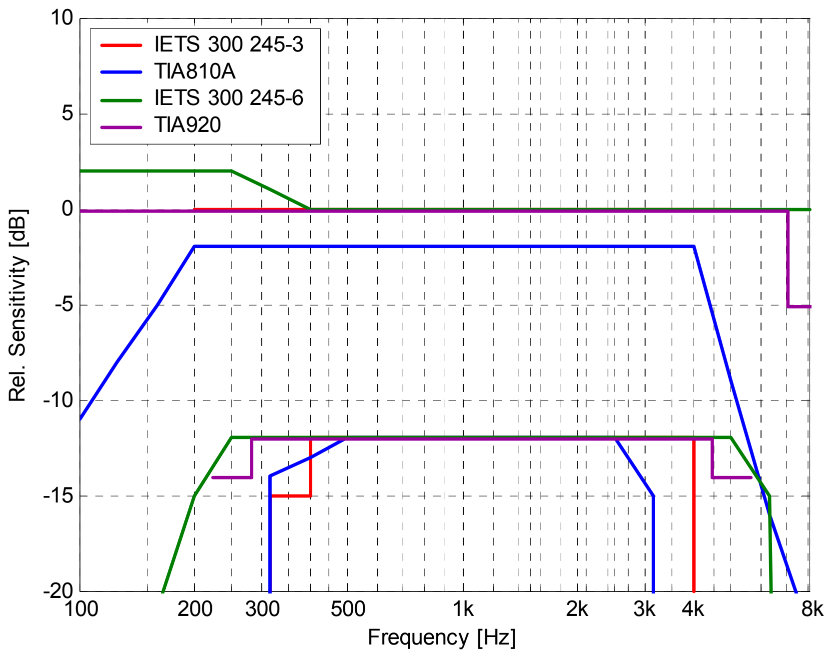

Figure 4: Hands free receiving frequency response masks

The main difference between narrowband and wideband requirements is the extended range in the high frequency region and, even more important, in the low frequency region. While narrowband handset telephony has a minimum frequency range of approximately 300 to 3.4 kHz, wideband telephony requires roughly a range from 160 Hz to 6.3 kHz.
But also between the frequency masks of the different wideband standards, there are significant differences. Both for handset and hands free, the European recommendations [2][3] demand a broader bandwidth than the US recommendation TIA 920. The TIA 902 also allows for an emphasis of high frequencies in the handset receive direction, which is not allowed by other standards

Artificial Ear

For narrowband handset measurements the TBR8 specifies the use of a type 1 artificial ear with an option to use a 3.2 low leakage ear. The TIA 810A only states to use a "suitable" ear type.
For wideband measurements, the I-ETS 300 245-5 specifies the use of a type 3.2 low leakage ear or, where not possible due to the geometry of the handset, the use of a 3.3 ear. TIA 920 only states not to use a type 1 ear, but gives no specific guidance on the ear type to use. For both wideband standards, the measurement is made at the DRP (drum reference point) and the result has to be corrected to the ERP (ear reference point).
The main difference between the measurement with a type 1 and a type 3.2 ear is the difference in the low frequency range, when traditional handsets are measured. This deviation, known as the "leakage effect", is caused by the leakage of the 3.2 artificial ear. Figure 5 shows the measurement of a conventional handset with a type 1 and a type 3.2 low leakage artificial ear. With a type 3.2 ll ear, the level is approximately 10 (15) dB below the level measured with a type 1 ear at 150 (100) Hz. For narrowband telephony this effect is not very strong, but for wideband, the leakage effect leads to a significant difference in the perceived signal, when the handset is held sealed or loosely to the ear. A high leakage sensitivity leads either to a signal that has too much emphasis on the low frequencies (sealed condition) or a signal, which has too much emphasis on the high frequencies (leakage condition). A balanced sound impression in both listening conditions can only be achieved with a low leakage sensitivity.
The most realistic measurement set up is probably the measurement with a type 3.4 ear. But this measurement has a limited reproducibility due to the dependence on the contact pressure. The type 1 ear is not the most realistic condition, but reflects a situation, where the user holds the handset sealed

to the ear. Especially when it comes to evaluating the maximum sound pressure level (acoustic shock measurement), a measurement with a type 1 ear has to be carried out.
Regarding reproducibility and realistic use cases, the type 3.2 ll ear seems to be a good compromise. Therefore measurements with type 1 and type 3.2 ll ears should be made to evaluate the leakage sensitivity and make sure, that the acoustic shock limit is not exceeded under all listening conditions.

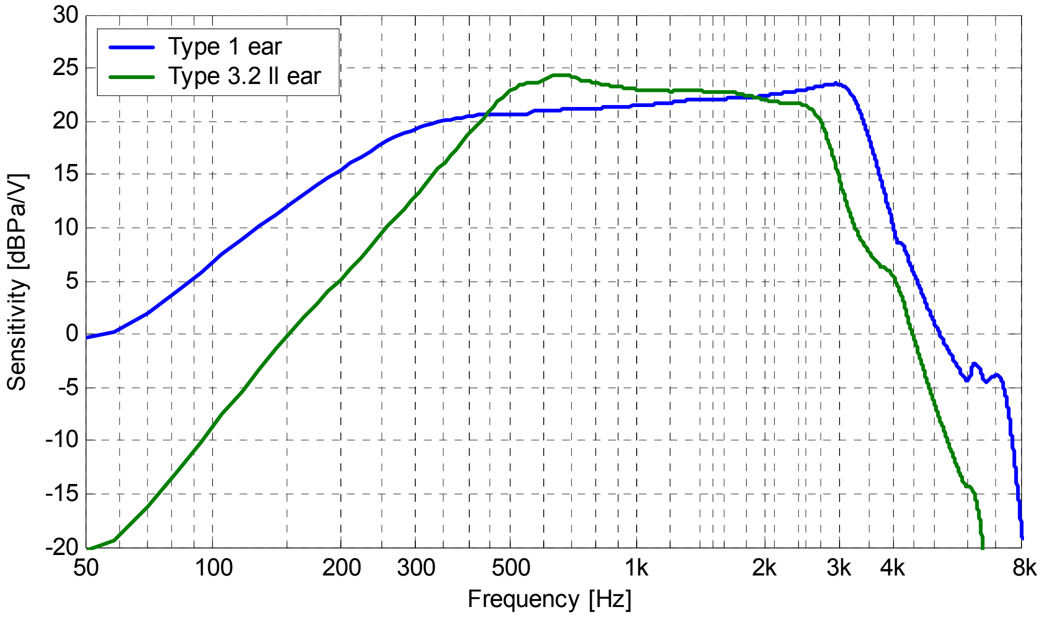

Figure 5: Receive frequency response of a conventional handset measured at a type 1 and type 3.2 ll artificial ear

Suitability of conventional transducers for wideband telephony

Besides meeting the electroacoustical requirements posed by the standards, the transducers in telephones have to fulfil additional requirements, mainly size limitations and cost. Commercially available transducers are tuned to conventional narrowband telephony requirements and have a limited capability to meet wideband requirements.

Microphones

The usually employed electret condenser microphones have by nature a flat frequency response over a wide frequency range. The only problem, which can arise, is caused by mounting the microphones with a small opening and a large front volume in the plastics (e.g. to improve EMI compliance). Such a mounting causes a Helmholtz resonator, which cuts off the high frequencies.

Receivers

Figure 6 shows the response of a typical receiver used for telephone handsets, measured with a type 1 ear. The frequency response is dominated by three resonances and is tuned to meet the narrowband frequency response requirements without additional electrical analogue or digital filters. Such a component cannot easily be modified to meet the wideband requirements. Moving the resonances to cover a broader frequency range, will lead to a bumpy frequency response, which cannot easily be equalized.
Conventional receivers also have a high mechanical impedance, which results in a high leakage sensitivity of the frequency response. To achieve a lower leakage sensitivity a low impedance transducer has to be employed and the receiver has to be mounted with a leakage in front of the receiver inside of the handset. Such a leakage can cause additional problems like acoustic coupling

between the receiver and the microphone or acoustic coupling between receiver front and back volume if the receiver is no closed at the backside.
Additional requirements for handset receivers are:

- robustness (a handset is often dropped)
- integration of an additional coil to support hearing aid coupling
- low cost
- a sealed back volume to minimize acoustic coupling to the handset microphone (TCLw requirements)

While extending the high frequency range of conventional receivers poses not a major problem, extending the low frequency range is. The closed back volume of many receivers limits the low frequency range and the high mechanical impedance increases the leakage sensitivity.
Extending only the high frequency range leads to an unbalanced listening impression. The signal sounds nervous and unpleasant. As a rule of thumb, the signal range should be extended similar to the low end and to the high end from 1 kHz on a logarithmic frequency scale.

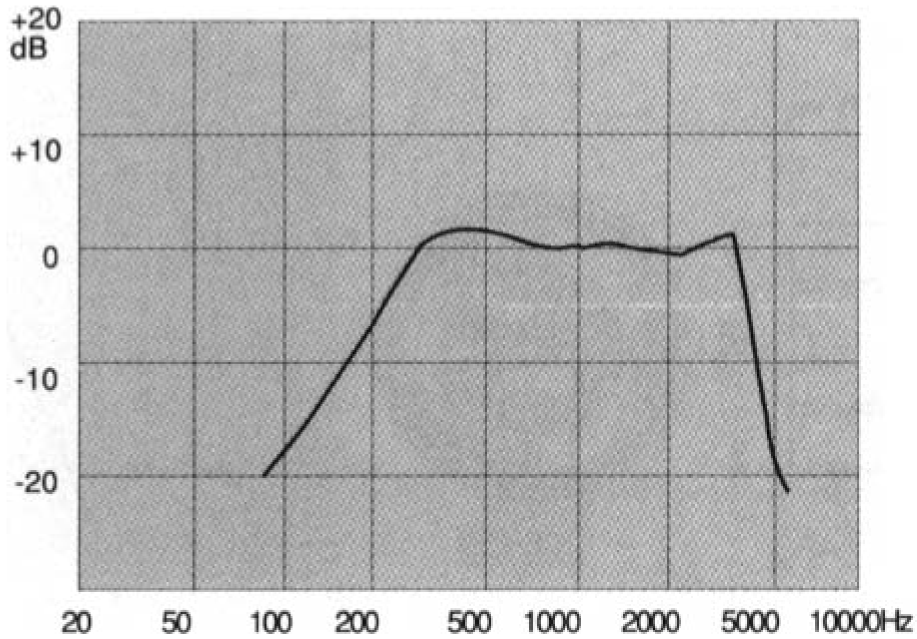

Figure 6: Typical handset receiver response measured at a type 1 ear

Speakers

Figure 7 shows the typical frequency response of a 50 mm diameter loudspeaker. The resonance frequency of such speakers usually lies above 300 Hz. When using smaller speakers, e.g. for mobile terminals, the resonance frequency is even higher. The resonance frequency is mainly dominated by the suspension of the membrane and the size of the speaker (weight of the membrane and coil). Since all components for consumer terminals are very cost sensitive, the suspension of such speakers is usually made out of the same material as the membrane itself. Since the membrane itself has to be made out of a stiff material, the suspension is also relatively stiff. Mounting the speakers in a housing with a small back volume usually additionally increases the resonance frequency. Another problem is the high Q factor of such speakers, which is caused by a small magnet. All these factors limit the low frequency response of such speakers.
In the high frequency range, conventional speakers usually transmit enough energy, though not with a flat frequency response. Like discussed in the section about receivers, extending only the high frequency leads to an unpleasant listening impression.

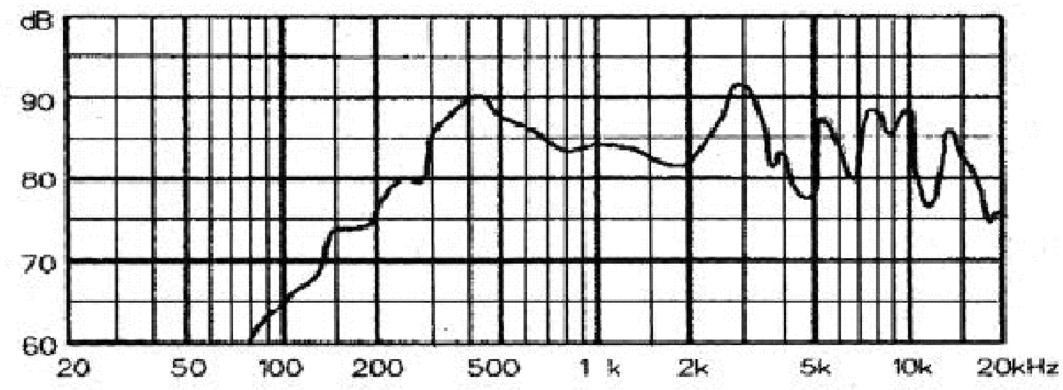

Figure 7: Typical speaker response of a 50mm diameter speaker (measured in baffle)

A speaker that is more apt for wideband telephony has to have a lower resonance frequency. That means larger size, more elaborate suspension and a bigger magnet. All these factors will result in higher cost.

Consequences for the E-Model and “objective” speech quality measurements

As discussed in the preceding paragraphs, conventional transducers will fail to meet all requirements of the wideband standards. This is especially true for mobile terminals. But even without meeting all the requirements, an extended frequency range could be transmitted, resulting in an improved speech quality and intelligibility.
Therefore several degrees of wideband should be defined (e.g. type 1,2 and 3) to characterize how far the frequency range is extended from narrowband to wideband for a certain system. These degrees should be characterized with a certain quality value resulting in a higher MOS score.
Care has to be taken to make sure that the frequency range is not only extended into one direction (high frequencies), but that a balanced frequency response is maintained. An unbalanced frequency response has to lead to a lower score than a balanced frequency response.
Also the leakage sensitivity of a system has to be rated. A high leakage sensitivity has to lead to a lower score than a system with a low leakage sensitivity.
Currently the E-Model considers only loudness ratings, not the exact frequency response. To evaluate wideband systems correctly, the frequency range, the spectral balance and the leakage sensitivity have to be considered too. The same holds for other "objective" speech quality measures.
Of course careful subjective listening test have to be carried out to find the right influence variables and to rate them correctly.

Conclusion

With the proliferation of IP-telephone terminals, both PC and conventional telephones, wideband speech transmission comes back into the focus.
While microphones used for standard narrowband terminals are well capable to transmit the extended frequency range, receivers and speakers aren't. But even with conventional transducers, an improvement in quality and intelligibility can be achieved though the frequency masks posed by the wideband telephony standards are not fully met. New components will provide better suitability for wideband speech transmission but in some applications, like mobile terminals, the low frequency range will never be transmitted fully.
The E-model and objective speech quality measurement algorithms should honour also limited extensions of the narrowband telephony frequency range which do not meet the wideband

requirements in total. But they also have to make sure, that spectral balance and leakage sensitivity are rated correctly.

Literature

[1] "I-ETS 300 245-3: Integrated Services Digital Network (ISDN); Technical characteristics of telephony terminals Part 3: Pulse Code Modulation (PCM) A-law, loudspeaking and handsfree telephony", February 1995
[2] "I-ETS 300 245-5: Integrated Services Digital Network (ISDN); Technical characteristics for telephony terminals; Part 5: Wideband (7 kHz) handset telephony", ETSI, December 1995
[3] "I-ETS 300 245-6: Integrated Services Digital Network (ISDN); Technical characteristics of telephony terminals; Part 6: Wideband (7 kHz), loudspeaking and hands free telephony", ETSI, January 1996
[4] "TBR 8: Integrated Services Digital Network (ISDN); Telephony 3,1 kHz teleservice; Attachment requirements for handset terminals", ETSI, October 1998
[5] "TIA/EIA-810-A: Transmission Requirements for Narrowband Voice over IP and Voice over PCM Digital Wireline Telephones", Telecommunication Industry Association, December 2000
[6] "TIA-920: Telecommunications; Telephone Terminal Equipment; Transmission Requirements for Wideband Digital Wireline Telephones", Telecommunication Industry Association, December 2002
[7] G.107: "The E-model, a computational model for use in transmission planning", March 2003, ITU
[8] http://www.globalipsound.com/pdf/gips_iSAC.pdf
[9] http://www.siemens.com/Daten/siecom/HQ/ICN/Internet/Enterprise_Networks/WORKAREA/skuch_c/templatedata/English/document/binary/a31002-h1000-a500-1-29_1156691.pdf
[10] P.57: "TELEPHONE TRANSMISSION QUALITY; Objective measuring apparatus; Artificial ears", July 2002, ITU