Wideband speech transmission is one of the technologies discussed for use in wireless systems and VoIP scenarios to provide a quality dimension significantly different from today’s narrowband communication systems. While the problem of wideband speech coding is mostly solved the requirements for a wideband speech service are still under discussion. 3G mobile networks and VoIP networks both allow the transmission of wideband speech. In order to take advantage of the high quality networks appropriate services and terminal solutions are required.

The paper gives an overview about the requirements in wideband communication systems. The parameters influencing the speech quality and the acoustical user interface and its influence on the perceived speech sound quality are introduced. Besides the speech sound quality and its perception by the user the requirements for a wideband speech service with respect to other performance parameters relevant for a communication service are discussed: performance of the (terminal) equipment under background noise conditions, echo cancelling characteristics, double talk capability and others.

1 Introduction

The telecommunications world is moving away from the traditional PSTN type networks to new IP based and wireless networks. Both techniques are no longer restricted to their 4 kHz bandwidth as used in the traditional telephone systems. In order to improve speech quality for the users of the new networks, wideband speech transmission is certainly one of the most interesting techniques which might help to increase the acceptance and usage of the new types of networks.

While wideband speech coding techniques have been developed and standardized over the last years with quite good success (see [1], [2]) the speech quality parameters influencing the overall perception of speech communication in wideband scenarios have not been investigated thoroughly. Speech quality assessment and prediction as well as planning tools for wideband systems also are not available yet. The parameters determining the speech quality of wideband systems and their individual contribution to the quality perceived by the user have not yet been fully understood. Therefore neither the tools for speech quality assessment in wideband systems nor the tools for modeling and speech quality prediction of wideband systems complete. The different parameters influencing the speech quality, the methods and tools available up to now are described in general in the following sections.

2 Parameters Influencing Wideband Speech Quality

In a first step, the parameters influencing the wideband communicational quality can be regarded similar to the ones which are relevant for narrowband speech transmission. Fig. 1 gives an overview about these parameters and how they translate into wideband scenarios.

![Figure 1: Speech quality parameter](image)

Although the general parameters contributing to the speech quality are the same, the overall speech quality
perception is different for wideband systems compared to narrowband communication systems. While the influence of wideband transmission on the speech sound quality and on the perceived loudness of the speech is investigated and understood rather well [3, 4] the influence of wideband transmission on other parameters is unknown. It can be expected that in a first approximation the influence of delay and echo should be similar to the influence known from narrowband transmission systems. However, specifically the frequency characteristics of echo signals might have a bigger influence on the perception of the echo signal compared to the narrowband case. The influence of delay on the conversational dynamics as well as on the general perception of echo should be comparable to narrowband. The system performance with respect to double talk capability is in a first step not influenced by the transmission bandwidth, since this parameter is mainly determined by switching time between sending and receiving, the attenuation inserted during double talk and the echo loss during double talk.

One of the probably most important parameters different between narrowband and wideband systems is the quality of background noise transmission. The quality of background noise transmission is determined by the parameters quality of the transmitted background noise, speech sound quality in the presence of noise and background noise transmission quality while speech is present either from the near end or the far end subscriber.

### 2.1 Overall Speech Quality

Investigations reported in [4] show that in general the overall speech quality can be improved when using wideband communication systems, but only in special conditions. The investigation results are based on conversational tests. The tests reported in [4] show that for clean conditions (no packet loss, no background noise) the overall speech quality of wideband communication system is not improved significantly (see Fig. 2). The tests however showed that an almost systematic improvement is achieved for wideband and G.711 speech codecs in noisy or packet loss conditions. It should be noted that these results are not necessarily in agreement with other investigations. It can be assumed that due to the conversational test design and the variety of quality parameters asked to the subjects the test subjects might have had difficulties in assessing the different quality parameters and separate them perceptually.

![Figure 2: Global quality of wideband communication from [4]](image)

### 2.2 Intelligibility

Figure 3 shows the results of tests conducted in [17] for comparing the speech intelligibility between wideband systems and narrowband systems.

![Figure 3: Intelligibility [17]](image)
Logatome tests were used in order to get a better discrimination between the (small) differences in speech intelligibility between narrowband and wideband systems. The result indicates that speech intelligibility increases by about 15% especially under noisy conditions. This increase of intelligibility is certainly remarkable.

### 2.3 Speech Sound Quality

A variety of subjective tests have been conducted to evaluate the increase of the subjectively perceived sound quality. Results reported so far depend on the type of tests used. While [4] reports no significant difference between narrowband and wideband communication systems with respect to the parameter sound quality, the results published in [7] and [8] indicate an improvement in sound quality. In the experiments reported in [4] conversational tests were used and a variety of parameters had been asked to the subjects. In [7] and [8] third party listening tests had been used and only the parameter speech sound quality was tested. [7] shows an MOS offset (of about 0.5) between narrowband and wideband systems. Other experiments [8] indicate that the “quality difference” between narrowband and wideband signals may also be in the range of an AMOS 1.0 to 1.3 depending on the way of presenting the signals.

A general problem for wideband systems is the use of a traditional handset. Due to the limited space for acoustical transducers needed for presenting the speech signal in receiving direction the sound quality of such transducers is typically poor. The main reason is the absence of low frequency components. Due to the small dimensions of the speakers the acoustical source impedance below about 1000 Hz is high. The acoustical leakage between handset and human ear does not allow an acoustically close coupling of the handset leading to a low load impedance and resulting in a decrease of the low frequency components. When using headsets this situation can be improved greatly since the headsets can be coupled more tightly to the ear. Well designed headsets do not show any decrease in sound pressure level in the low frequency domain and therefore seem to be ideally suitable for wideband communication, especially in mobile applications.

The use of hands-free systems seems to be the preferred technique for wideband applications and may lead to a superior speech sound quality if well designed loudspeakers are used. This is possible for all types of fixed terminals however loudspeakers have to be improved significantly compared to the existing narrowband terminals. Mobile terminals will suffer from the limitation given by the speaker size and the limited power consumption.

### 2.4 Quality of background noise transmission

Background noise may be present during the whole entire call. Spectral content, temporal behavior and the signal to noise ratio between the background noise signal and the transmitted speech signal may vary highly depending on the communicational situation and the environment the user is conducting his call. From the perceptual point of view the quality of background noise transmission has to be described by three parameters:

- the transmission quality of background noise while no speech is present
- the quality of the transmitted speech while background noise is present simultaneously
- the quality of the transmitted background noise signal while far end speech is present

Although this general description is applicable for narrowband systems as well, masking effects (e.g. for low frequency background noise signals such as found in cars) are expected to have a different influence on the perceived speech quality. [9] describes the test setup for conducting subjective speech quality tests in the presence of background noise. Currently no test results are available for wideband systems based on the test procedure described in [9].

### 3 Test Procedures

#### 3.1 Acoustical Interface

For wideband systems the acoustical interfacing of terminals as used in the past [10] are no longer applicable. The most appropriate acoustical access interface is the use of a Head and Torso simulator (HATS) as described in [11]. The HATS can be used for handset terminals, headset terminals, but for all types of hands-free terminals as well. The electrical interfacing of wideband systems in general is not different to narrowband speech communication systems. The use of HATS also is the only way to provide a common acoustical interface for devices developed for the transmission of high quality audio (high quality headphones) and acoustical interfaces developed for communicational use.

#### 3.2 One Way Speech Transmission

The assessment of speech quality in single talk situations is well established. For the measurement of the speech loudness the wideband loudness ratings as defined in [3] are applicable. Frequency response characteristic requirements are found already in various
standards. Ideally the receiving response characteristics should be close to the average free-field characteristics of humans. Also all the requirements for loudness ratings should be transferred to the free-field reference point, however – due to the historical use of the ERP (Ear Reference Point) the transformation of the free-field characteristics to the ERP seems to be problematic. In the long term the reference systems in telephonometry should be transferred to the free-field reference which is commonly used for high quality audio applications.

In principle the speech sound quality in wideband systems can be evaluated using reference signal based and psychoacoustically motivated testing techniques described e.g. in [12] and [13]. While the procedure known as “TOSQA2001” [13] has been partly validated for wideband applications and acoustical interfaces the “Wideband PESQ” method as described in [14] may be applicable for wideband electrical to electrical connections, but not including the acoustical interface. None of the methods has been standardized yet.

The basic relationship between the perceived speech sound quality in wideband scenarios compared to the well known narrowband scenarios is described e.g. in [17] (see Fig. 5). Based on such data the wideband extension of the psychoacoustically motivated test procedures can be made.

3.3 Double Talk Performance and Switching Characteristics

An important factor in communicational systems is their conversational performance. The double talk capability including all types of switching have an important influence on the communicational quality. The testing of double talk and switching characteristics is well-established and well-linked to the results of subjective testing for narrowband systems (see [15], [16]). The testing procedures used are the combination of speech-like signals either combined in time or in frequency. When the test signals are combined in time, typical composite source type signals [15] are used. The combination of speech-like signals in frequency are typically based on the use of voiced sound simulating signals inserted simultaneously in both transmission directions.

In a first approach there is no evidence that such parameters are perceptional different in wideband systems, however, the validation of this assumption has not yet been made.

3.4 Background Noise Transmission

For testing the proper simulation of background noises in a lab type environment is required. Newer investigations as described in [17] show that a “4.1 loudspeaker arrangement” used for the playback of binaurally recorded sounds allows the simulation of pre-recorded background noises in a very realistic manner. Such an arrangement (see Fig. 5) can be used for all sorts of terminals.

Figure 4: Relationship between narrowband MOS and wideband MOS (from [7])

Figure 5: Test setup for background noise simulation in lab-type environments [17]

The analysis methods - which could be used for background noise and speech in the presence of background noise – are still under investigations.

The evaluation of the pure quality of background noise transmission can be made using the method “Relative Approach” [18]. This method takes into account the
human perception properties with respect to short and long-term signals as well as sensitivity of the human ear against dominant structures in frequency. Aurally obvious structures such as instantaneous events in time or dominant spectral structures are displayed on a time/bark scaled “spectrogram”. The method (Fig. 6) can be applied to stationary signals without using any reference signal.

First experimental results are discussed in [19]. For the tests of a variety of wideband speech codecs a specific background noise signal simulating calls of organ music has been developed. Figs. 7 displays the three-dimensional result of the method applied to a G. 722.1 wideband codec [20] in comparison to the AMR wideband codec [1]. It is obvious that the AMR introduces much more artifacts to this test signal compared to the G. 722.1 wideband codec. This result from the objective analysis is in line with the results of the subjective listening tests (reported in [19]). Further work is required to validate the method with a variety of impairments and to map the 3-dimensional result on a 1 dimensional quality value.

4 Modeling Wideband Transmission

The world-wide accepted model for transmission planning is the E-model as described in [21]. This model covers the “traditional” parameters of narrowband communication systems. The current model allows to some extend the prediction of the transmission rating for non-linear and time-invariant devices such as speech coders. However, due to its simple construction and the assumption of the independency and additivity of all the parameters on a psychological scale more complex and time-variant behavior of transmission systems such as the influence of double talk, the various types of switching characteristics, the different implementations of speech echo cancellers can not be modeled.

The current model can not be used for wideband transmission. Proposals have been made by [22] and [8] – based on subjective test results for overall speech quality and wideband listening speech quality to extend the R-scale which is the basis for the transmission rating. The general principle followed for the extension is to compare the result of mixed subjective tests (narrowband and wideband) to the results of narrowband only subjective tests. These results indicate a relationship between the “narrowband mean opinion score” and a “wideband mean opinion score” (see Fig. 5). Based on various subjective results [7] and [8] propose the extension of the model. More details can be found in [22].

5 Summary and Outlook

The determining quality parameters for wideband communication systems are known in general, however for various parameters contributing to the perceived overall speech quality basic subjective experiments are still missing.
Testing procedures for some parameters in wideband systems are available and briefly described, however, for one of the main influencing parameters namely the speech quality in the presence of background noise and the quality of background noise transmission no testing techniques are available. Further work is required to develop the subjective basis for the development of the objective methods.

6 References

[1] 3GPP TS 26.171: AMR speech codec, wideband; General description, 3GPP 03/2001