Introduction of **Acoustic Testing** of mobile phones

Gaoxiong Yi

yigaoxiong@caict.ac.cn

China Academy of Information and Communication Technology

Course Objectives:

Bring a brief introduction of testing electro-acoustic performance of mobile phone to the audiences, focusing on

What to test
How to do the test and
Which standard should refer to

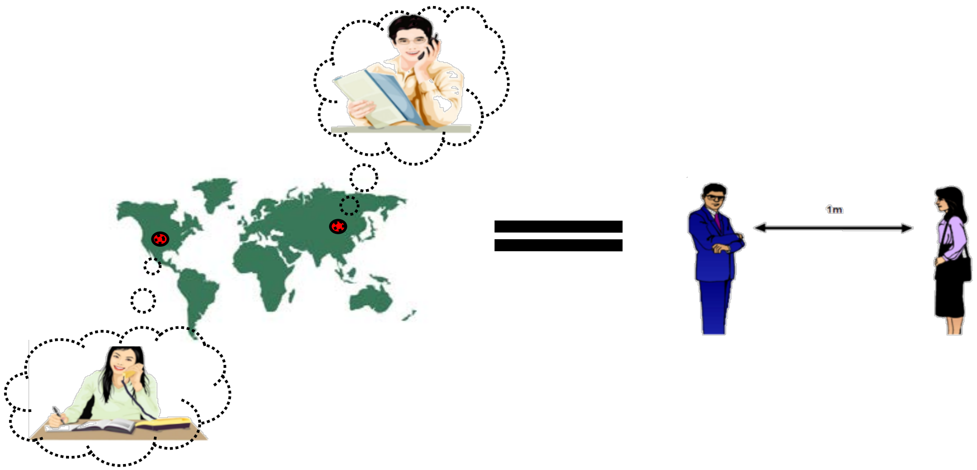
Outline

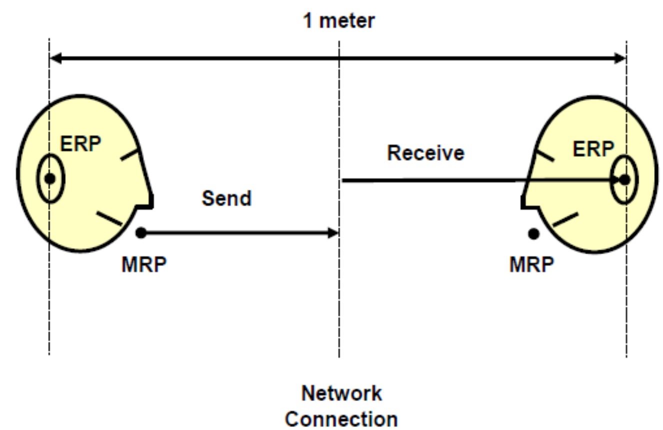
- The basic of electro-acoustic testing on telephone
- Test setup and test instruments
- Typical test cases of mobile phone
- Current standards in use for mobile terminals
- Future prospects

The target to design

What is the target acoustic (speech) performance of mobile terminals?

- The Ortho-Telephonic Reference





Simulating Acoustic path between a talker and a listener, facing each other at a distance of 1 meter in the free field

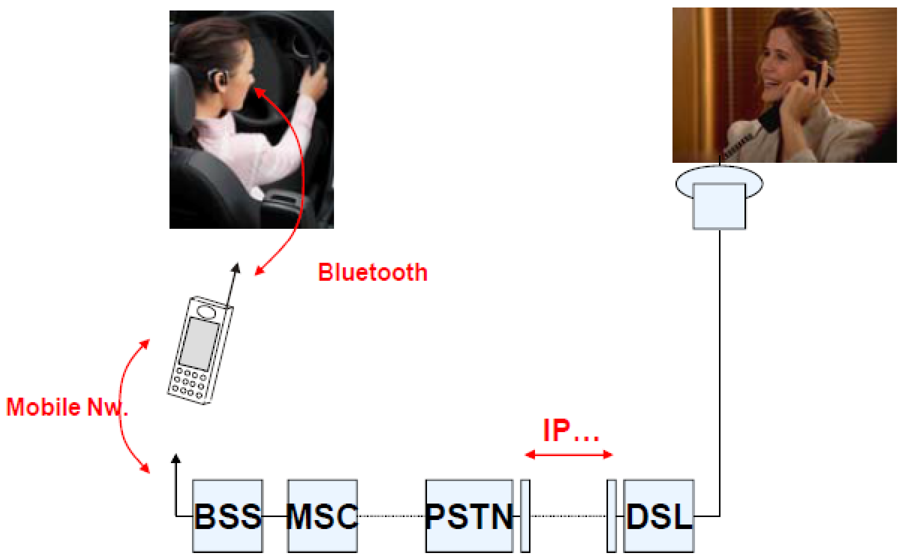
To inter-connect with the network, the acoustic path is divided into two parts:

Send: terminal to network

Receive: network to terminal

The inter-networking

What are the mobile terminals facing with when making speech communication



Various scenarios:

Mobile internal networking (GSM/CDMA/TD-SCDMA/WCDMA/LTE)

Mobile to PSTN

Mobile to cordless

Mobile to DECT

Mobile to Bluetooth

Mobile to VoIP

Mobile to Car-handsfree

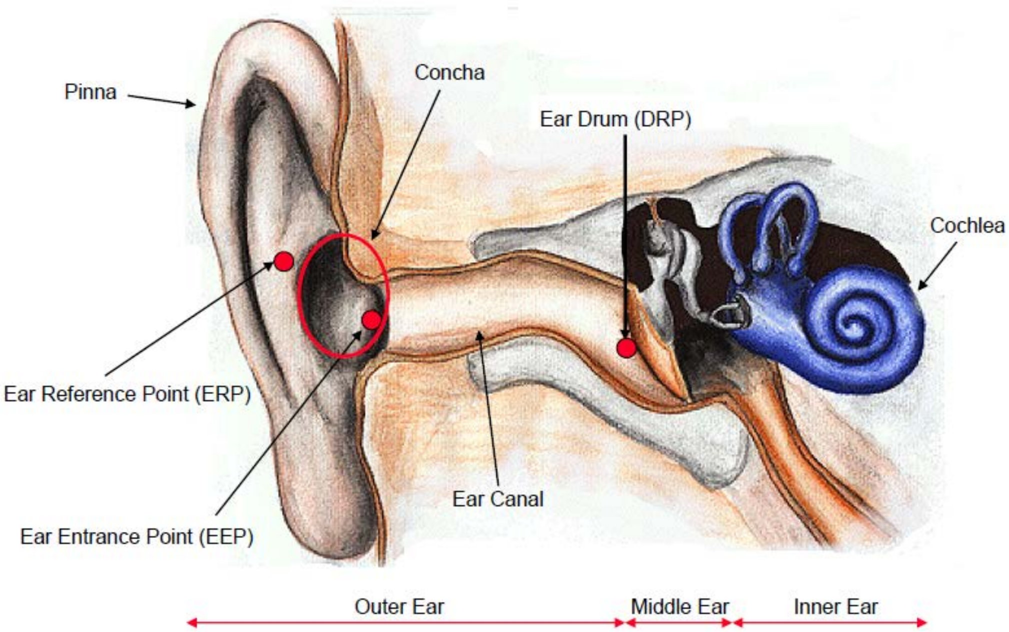
….

How the quality of speech communication could be guaranteed?

Stick to the design target no matter what the physical appearance and accessing technology is. Improve the speech performance as much as possible.

Human perception

How people hear sounds generated by mobile terminals



The outer ear :

consists of the pinna, the ear canal and the eardrum. It plays a role in the sound localization and transforming air vibration into eardrum vibration.

The middle ear:

Consists of the bones and muscles in the middle cavity.

It contract in response to loud sound, Transforming huge vibration changes at eardrum to tiny vibration changes at inner ear nonlinearly. This process helps to protect the sensory part of the inner ear from damage by loud sounds.

The inner ear:

Consists of cochlea , Transforming vibration from middle ear to 'waves' in cochlea fluid, so that can be detected by the hair cells , which turn 'waves' into electrical signals that travel up the auditory nerve to the brain.

How people perceive speech quality

The subjective opinion on speech quality is affected by a lot of factors, and may vary between people

The main objective of acoustic test on mobile phone:

To simulating certain situations and evaluate the acoustic performance, that might affect people perception and quality opinion, of the mobile phone accordingly, to guarantee certain user satisfaction for most of the people.

Outline

- The basic of electro-acoustic testing on telephone
- Test setup and test instruments
- Typical test cases of mobile phone
- Current standards in use for mobile terminals
- Future prospects

Devices used in test

The artificial ear used to capture acoustic signals generated by mobile terminals

- The artificial ear standardized in ITU-T P.57:

consists of the IEC 60318-4 occluded-ear simulator, to which is added an ear canal extension terminated with a pinna simulation device, as follows:

- Type 3.1 Concha bottom simulator.
- Type 3.2 Simplified pinna simulator.
- Type 3.3 Pinna simulator (anatomically shaped).
- Type 3.4 Pinna simulator (simplified).

type 3.2 Simplified pinna simulator. (high/low leak)

type 3.3 pinna simulator.

Type 3.3 artificial ear has been used most widely now

It provides not only similar shape but also similar acoustic impedance along with the ear canal simulator

Devices used in test

The artificial mouth used to generate acoustic signals

- The artificial mouth standardized in ITU-T P.51:

The artificial mouth need to keep good linearity, frequency response and have typical directivity to generate similar sound field around human mouth

Above: the artificial mouth including the lip-ring

Principle for calibration of the Artificial mouth (at MRP)

Devices used in test

150 ITU 1865-2015

CAICT 中国信息通信研究院 China Academy of Information and Communications Technology

Artificial “Head” with the artificial ear (P.57 type1/type3.2) and mouth (P.51) to place the Handset

Pressure Adjustment to Press phone Against ear

Artificial ear

Pliers

Artificial mouth

This “artificial head” includes the artificial ear and the artificial mouth.

The earpiece of the handset is placed over the artificial ear. The handset microphone is automatically placed in front of the Artificial mouth.

Devices used in test

The Head and Torso Simulator (HATS) defined in ITU-T P.58

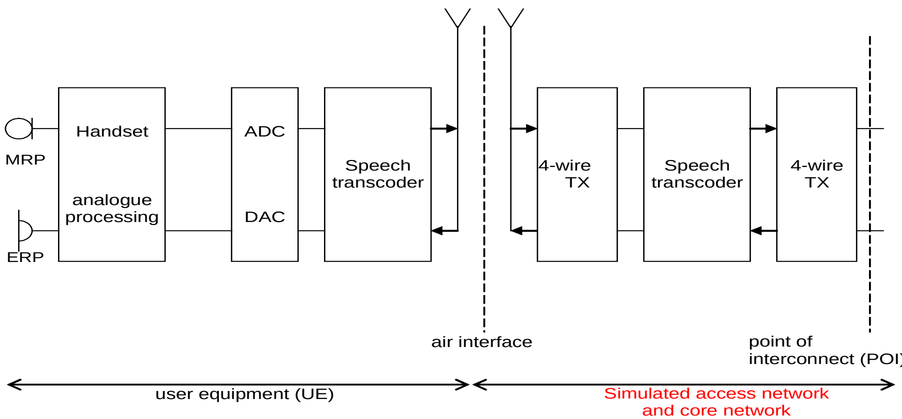
Head and Torso Simulator have built in artificial mouth and equipped with one or two type 3.3/3.4 artificial ears.

The improvements of the artificial ears and mouths, reproducing better the human ears and mouths behaviors. as a consequence, nowadays, the use of HATS is recommended for testing all types of speech terminals

Devices used in test

The electro-interface that sends signals to and captures the signals from mobile terminals

- The system simulator



The system simulator provides a access network and simulate the core network functions

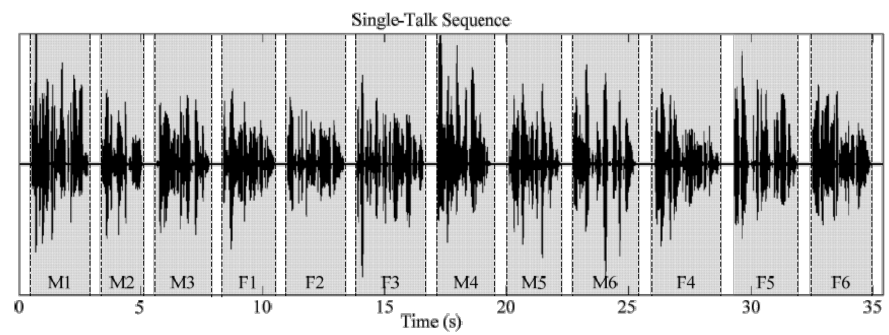
It allow a speech connection with mobile terminals. In addition to that, it contains a high quality reference codec with characteristics as close as possible to ideal, to convert the digital input/output bit-stream of the system simulator to the equivalent analogue values, for the purpose of input test stimuli to and obtain result speech signals from mobile terminals

Stimuli used in test

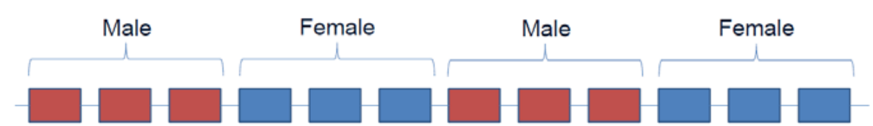
Various test signal are used, depending on the test purpose. At most cases, Real speech defined in P.501 are now used as test signal

The signal processing modules implemented in speech terminals are able to differentiate speech from other types of audio signals. Terminals will behave more natural and show performance close to real when facing real speech test signal in the tests.

A typical "single-talk" sequence of sentences spoken by all 12 speakers from the reference speech samples is shown.



The structure of the male/female single talk sequence

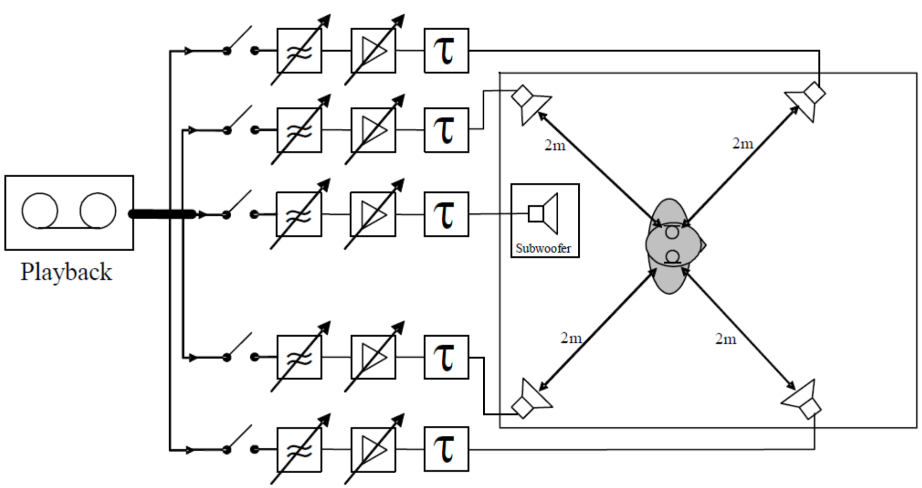


Speech material*:
- Psytechnics speech data base
- Nokia speech data base
- China CTTL speech database

Stimuli used in test

Background noise signals

Terminals have built-in processing functions to reduce the ambient background noise. When testing terminals in noisy scenarios, background noise signals are needed.

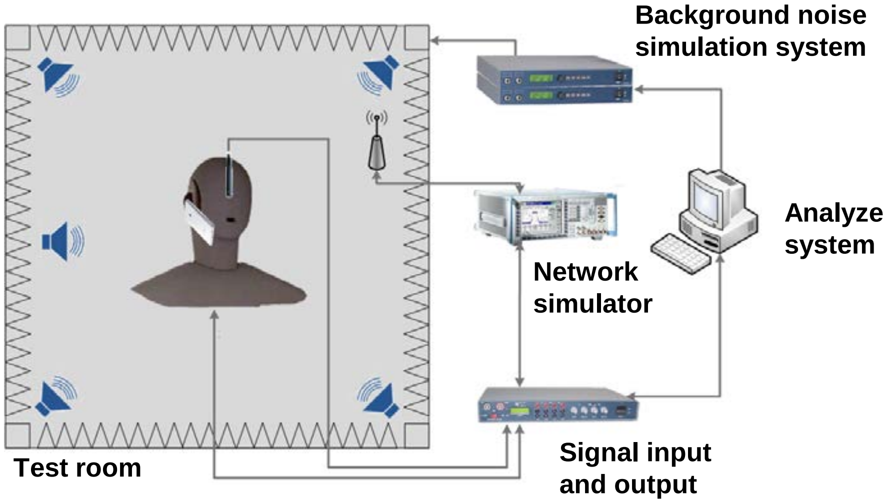


Background noise are simulated according to ETSI ES 202 396-1 standard.
To reproduce binaural recording, which was record with HATS equipped with both ears in chosen scenarios, and get similar noise level and frequency characteristics around the HATS placed in the test room.

Typical secnarios
Recording in pub
Recording at pavement
Recording at pavement
Recording at departure platform
Recording at the drivers position
Recording at sales counter
Recording in a cafeteria
Recording in business office

General test configuration

The overall test setup

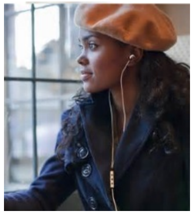


- Mobile Terminals are placed on the HATS.
- HATS is placed in the test room (usually anechoic room) and in the center of background noise playback system.
- Speech call is established by the mobile terminal and network simulator.
- HATS and Network simulator are connected to the signal input and output interface of Analyze system.
- The Analyze system output test stimuli and obtain subsequent result speech, do the test analyzing.

General test configuration

For mobile terminal, there are several modes need to be considered

- Operation mode
  - **handset, headset and hands-free mode.**

Every mode need to be taken into consideration to make sure it provides sufficient users satisfaction.

- Accessing mode
  - **Circuit Switch(GSM, WCDMA, CDMA, TD-SCDMA) and Packet Switch (VoLTE).**

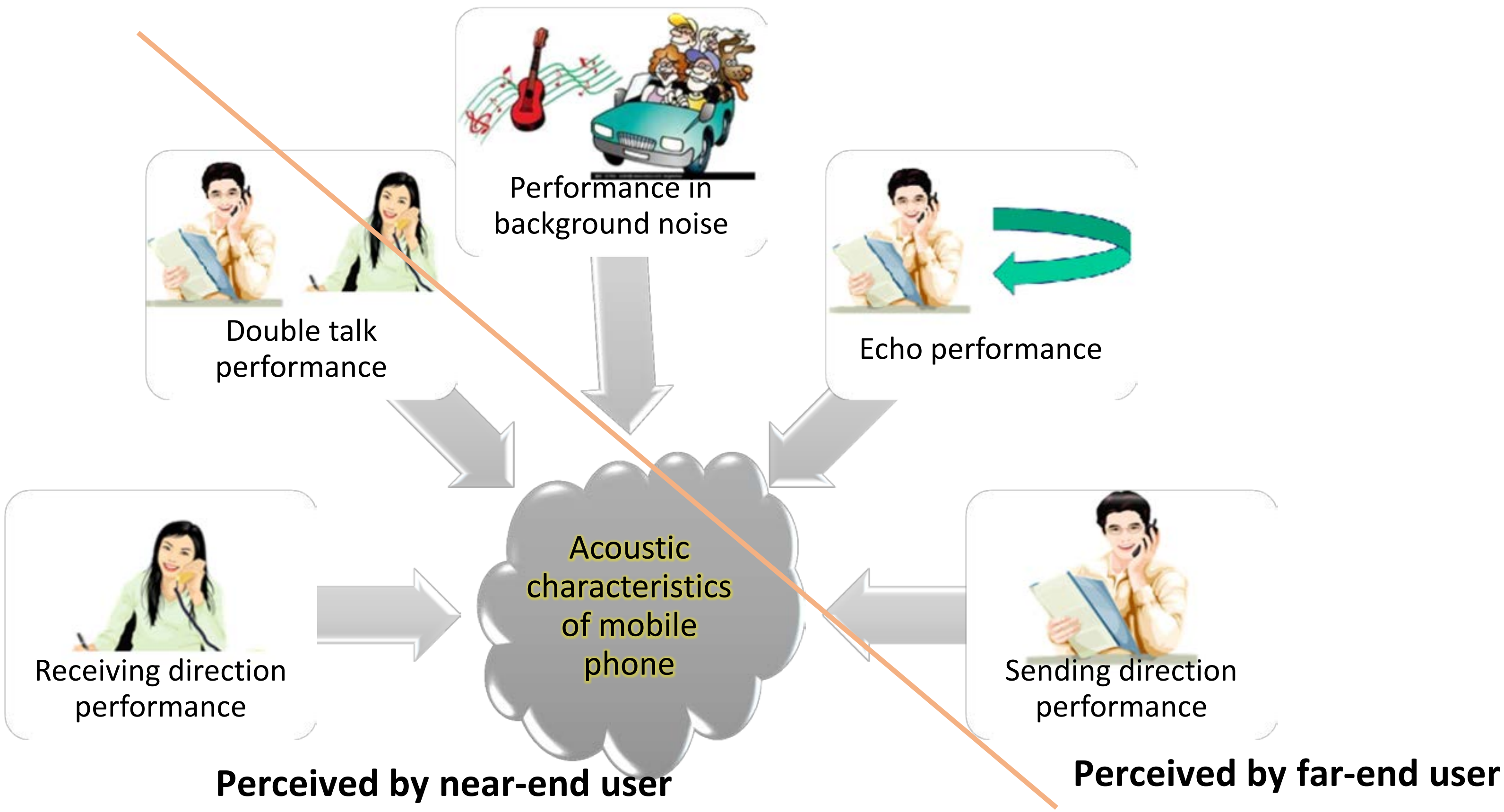
Many terminals will have different speech processing performance under different accessing mode due to its hardware and software design associated with the assessing technology it applied. So every mode should be covered.

Outline

- The basic of electro-acoustic testing on telephone
- Test setup and test instruments
- Typical test cases of mobile phone
- Current standards in use for mobile terminals
- Future prospects

Test case categorization

Generally the acoustic characteristics of mobile phone could be grouped into several category, depending on who it might impact and what it is dealing with



Basic test cases

Typically basic test cases are shown below:

Sending direction performance

Receiving direction performance

- Frequency Response
- Loudness rating
- Distortion
- Delay
- Idle channel noise
- Speech Quality Score
- SideTone

Double talk performance

- Double talk attenuation
- Double talk Echo performance

Echo performance

- Echo Loss
- Echo spectrum characteristics
- Echo temporal characteristics

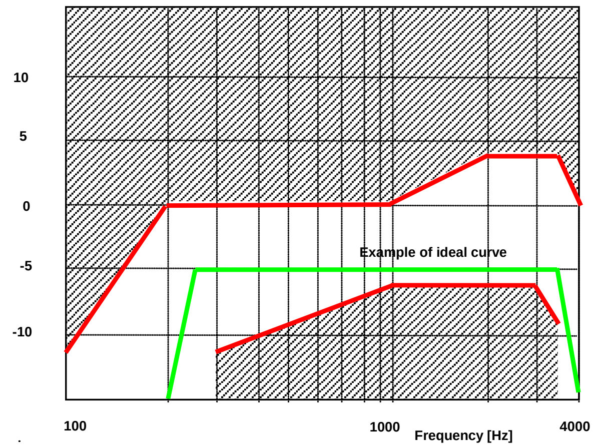
Performance in background noise

- Speech quality after noise reduction
- Noise quality after noise reduction
- Stability of noise transmission

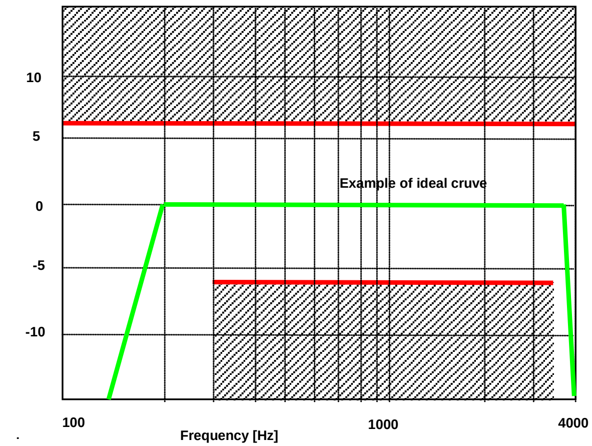
Basic test cases

Frequency response/ sensitivity

- Frequency response / sensitivity is used to evaluate how the frequency components are delivered from one user to another. Generally mobile terminals should not modify the speech energy distribution in frequency components too much so as to achieve basic "Fidelity" in the allowed frequency range and also guarantee certain speech intelligibility.
- The target overall frequency responses in send and receive is the ortho-telephonic reference response assumes an ideal frequency characteristic. The ideal frequency response is decomposed into frequency response in send an received respectively. Due to design constraints and inter-networking with traditional telephone terminals, allow tolerance were created.



Example Send Frequency response



Example Receive Frequency response

Basic test cases

Loudness Rating

- The Loudness Rating is used to evaluate how loud the speech signal delivered by the terminals will be perceived psycho-acoustically by the user. It is not to indicate absolute physical level attenuation of the speech transmitted, but the loudness in the human perception.
- The Loudness Rating is based on subjective model, with comparison of an unknown terminal (handset) to a reference handset (IRS-handset defined in ITU-T P.48. The Overall Loudness rating (Send plus Receive) of a mobile terminal should close to 10dB

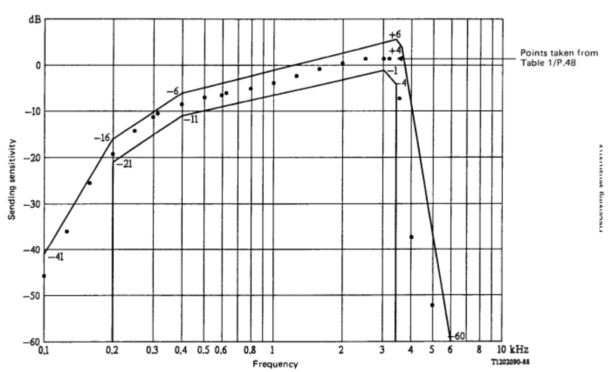


FIGURE 2/P.48

Suggested IRS sending mask

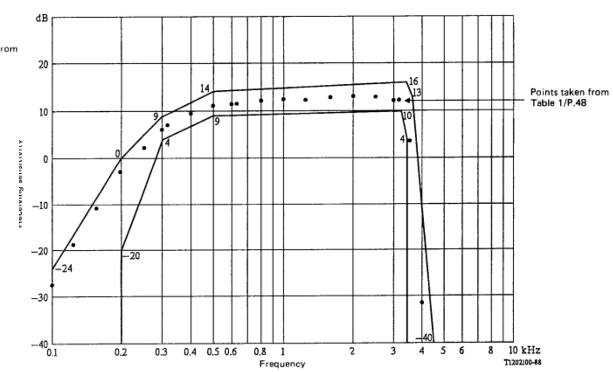


FIGURE 3/P.48

Suggested IRS receiving mask

Loudness rating is calculated based on the speech signal energy calculated in fixed bands and comparing these values to the loudness derived from the IRS reference system.

$$LR = -\frac{10}{m} \cdot \log_{10} \sum_{i=N_1}^{N_2} 10^{0.1 \cdot m(S_i - W_i)}$$

m = a constant (in the order of 0.2).

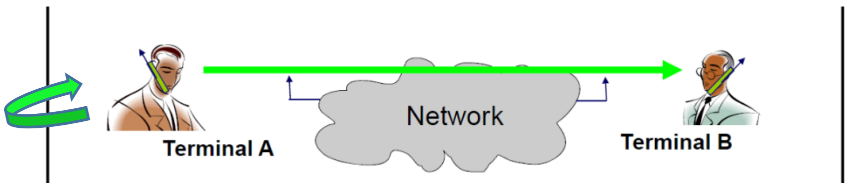
The summation is to be performed at frequencies F_i, spaced 1/3 octave apart.

W_i = weighting coefficient (different for the various LRs).

S_i = the sensitivity at frequency F_i of the electro-acoustic path under consideration.

Basic test cases

Sidetone (Sidetone Masking Rating)



When a phone user speaks, his own voice reaches his ear by several paths

- a) through the telephone side tone path;
- b) through the mechanical path within the human head;
- c) through the acoustic path to the ear and involving leakage at the earcap and human ear interface;
- d) through the mechanical path along a handset handle.

In general It is necessary to provide a terminal sidetone path for mobile terminals in handset and headset mode, Especially in case the human air-conducted sidetone paths are obstructed (one example being some binaural insert type headset UEs), to provide improve conversation naturalness and comfort.

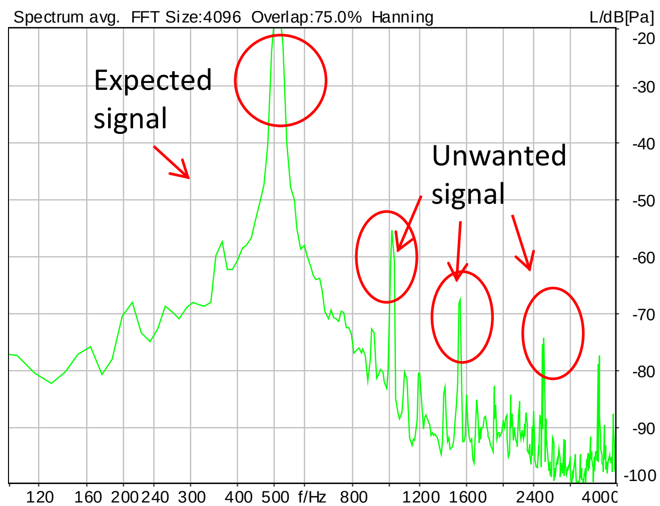
Sidetone masking rating is calculated comparing the generated sidetone by terminals and the user's own speech.

$$STMR = -\frac{10}{m}\log_{10}\sum_{M=1}^{N}10^{(m/10)(-L_{meST}-L_E-W_M)}$$

Basic test cases

Distortion

- **Distortion is used to evaluate whether or not the physical acoustic components of mobile terminals are working properly (components overloading, improper sound cavity design), and whether the signal processing in mobile terminals are not producing unwanted frequency components.**



Example when testing with pure tone signal 510Hz.
Harmonics such as 1020, 1560 and 2040 appears.

Distortion
Tested with pure tone of various levels and frequencies.

Sending level (dBPa at the MRP)	Sending Ratio (dB)
5	30
0	35
-4,7	35
-10	33
-15	30
-20	27

Frequency (Hz)	Receiving level at the digital interface (dBm0)	Receiving ratio at nominal volume setting (dB)
315	-16	20
408	-16	28
510	-16	28
816	-16	28
1 020	0	25,5
	-3	31,2
	-10	33,5
	-16	33,5
	-20	33
	-30	30,5

Basic test cases

Delay

- **Delay is the temporal interval introduced by the terminal and network between the time speech was produced at the mouth of near end user and perceived at the ear of far end user. It depends on the signal processing, coding, and accessing technology used by the terminals and networks.**
- **Delay is very important it has impact on echo performance and the dynamics and efficiency of voice conversation.**

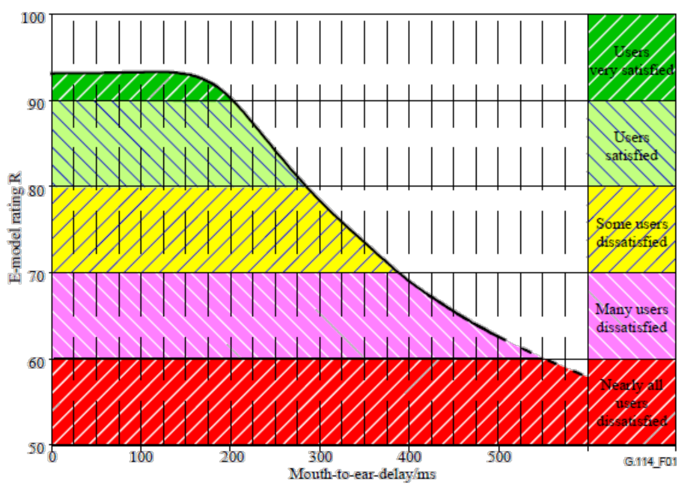
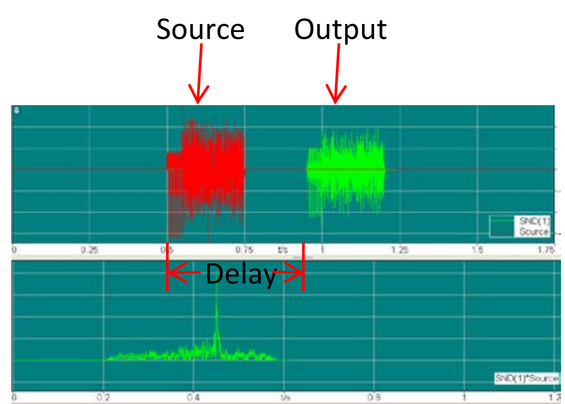


Figure 1/G.114 – Determination of the effects of absolute delay by the E-model

The conversation interactivity is considered to be transparent if delay is <150ms.

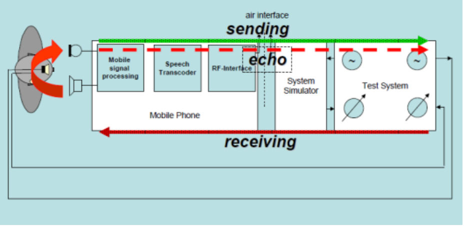
It is recommended to never exceed delay >400ms.
Considering the delay comprise both network delay and terminal delay, the delay of terminal, containing both in send an receive direction, should be less than around 200ms.



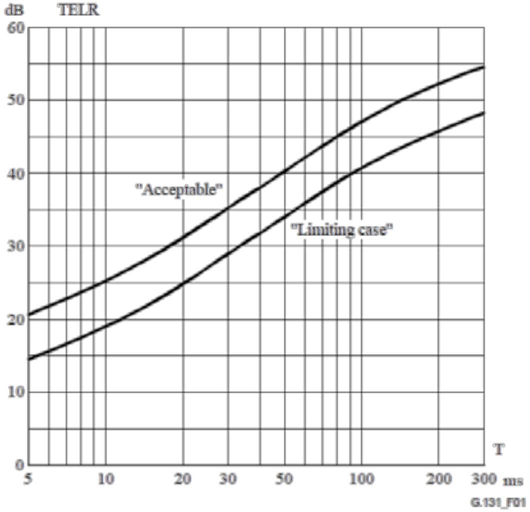
Basic test cases

Echo (Terminal Coupling Loss)

- **In most cases the echo user perceived is resulted by his/her own voice which transmitted to the far end user and then sent back by the far end microphone.**

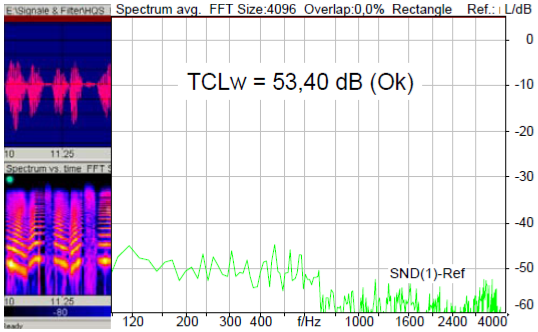


- **The degree of annoyance of talker echo depends both on the amount of delay and on the level difference between the original voice and the received echo signal. This level difference is characterized by the measure "Talker Echo Loudness Rating"**



T Mean one-way transmission time
TELR Talker Echo Loudness Rating

Main talker echo is depending on the ability to handle echo from the far end talker's terminal (Terminal Coupling Loss)



Basic test cases

Double Talk ability

- **Due to the Echo cancellation used in terminals, the duplex ability of terminal is affected and the conversation is not always fully transparent during double talk comparing to single talk.**
- **The most annoying effects during double talk, for users, are:**

  -- sentences, words, syllables interrupted or not transmitted completely
  during or shortly after/before double talk;
  -- echo during double talk.

The performance during double talk is mainly determined by two parameters: talker echo loudness rating and level variation between single and double talk (attenuation range).
In order to guarantee sufficient quality under double talk conditions, the TELR should be high and the attenuation inserted should be as low as possible.

Table 4/P.340

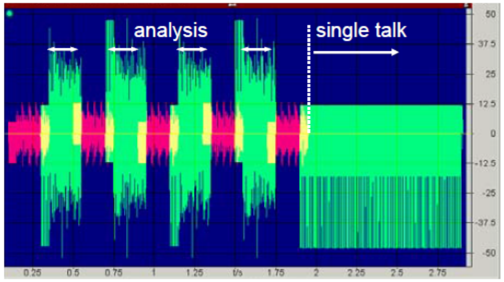
Sending direction	Receiving direction
Behaviour 1: $a_{H,S,DT} \leq 3$ dB	Behaviour 1: $a_{H,R,DT} \leq 3$ dB
Behaviour 2a: 3 dB $< a_{H,S,DT} \leq 6$ dB	Behaviour 2a: 3 dB $< a_{H,R,DT} \leq 5$ dB
Behaviour 2b: 6 dB $< a_{H,S,DT} \leq 9$ dB	Behaviour 2b: 5 dB $< a_{H,R,DT} \leq 8$ dB
Behaviour 2c: 9 dB $< a_{H,S,DT} \leq 12$ dB	Behaviour 2c: 8 dB $< a_{H,R,DT} \leq 10$ dB
Behaviour 3: $a_{H,S,DT} > 12$ dB	Behaviour 3: $a_{H,R,DT} > 10$ dB

Combination of two Composite Source Signals (CSS)

green: near end signal
red: far end signal

Signal description in ITU-T Recommendation P.501

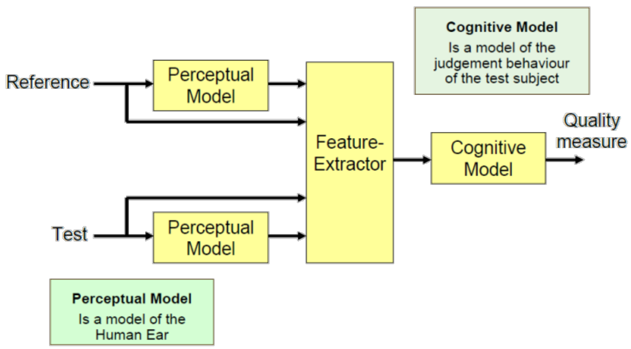
Description of analysis methods in ITU-T P.502



Basic test cases

Single talk speech quality

- **A score with 1-5 scale is sometimes desired to indicate the overall speech quality in single talk situation, simulating the human opinions on the single talk speech quality by taking into consideration those factors that might influence human perception.**



Impairment	Grade
Excellent	5
Good	4
Fair	3
Poor	2
Bad	1

- **There are a lot of speech quality accessing tools by modeling human auditory test results by comparing the processed signal and reference signal. And they evolves with the communication technology and scenarios.**

PSQM	PESQ	PSQM99	PAMS	TOSQA	PACE	VQI	POLQA
P.861	P.862	KPN	BT	T-Systems	Ascom	Ericsson	P.863

Currently the most advanced one is POLQA, ITU-T P.863 published around 2012.

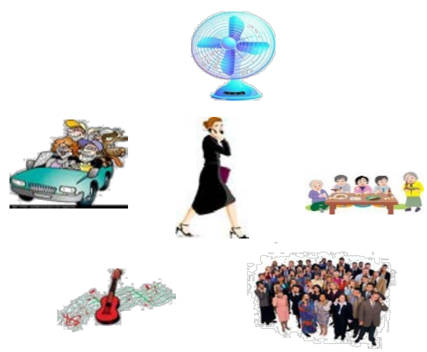
The speech quality score can be tested in send and receive. It is recommended to maintain score >3.5 to provide good quality. However, **the speech quality score is just one indicator of the acoustic performance of mobile terminals. It is not comprehensive enough.**

Basic test cases

Noise reduction quality

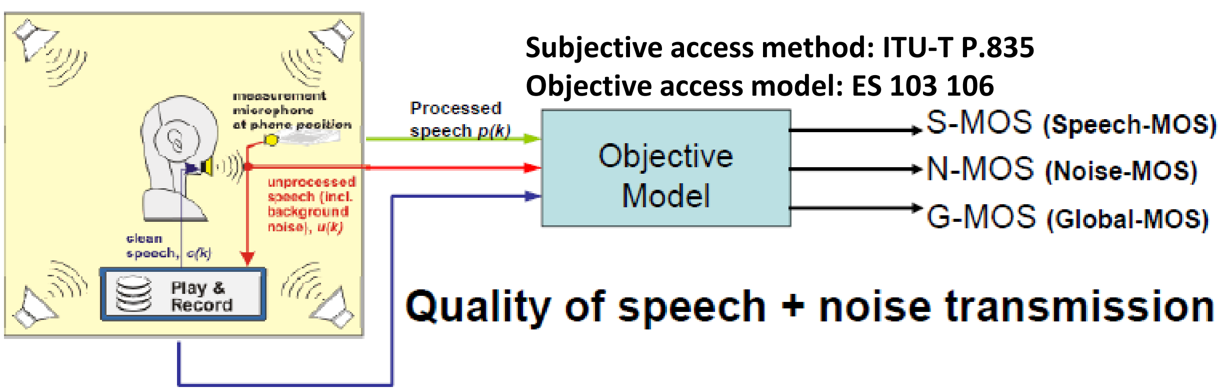
- **Similar to single talk speech quality, Scores with 1-5 scale is sometimes desired to indicate the performance of noise reduction algorithms used in terminals.**

Impairment	Grade
Excellent	5
Good	4
Fair	3
Poor	2
Bad	1



- In noisy environments, with the presence of noise, mobile terminals should at one hand deliver the speech as clean as possible, maintain the speech quality as much as possible. On the other hand, should reduce the background noise as much as possible, while deliver the residual noise as not intrusive as possible.

Subjective access method: ITU-T P.835
Objective access model: ES 103 106



It is recommended:
S-MOS >3.5
N-MOS >3.0

Basic test cases

Other test cases

- Idle channel noise
- Out of band signals
- Stability loss
- Maximum acoustic levels
- intelligibility assessments
- ……

Outline

- The basic of electro-acoustic testing on telephone
- Test setup and test instruments
- Typical test cases of mobile phone
- Current standards in use for mobile terminals
- Future prospects

Related Standards

There are many overall acoustic standards for mobile terminals specified in different levels

- International standards (ITU, etc)
  - **ITU-T P.313 Transmission characteristics for cordless and mobile digital terminals**

    This Recommendation is the general performance guidance for mobile speech terminals
- Regional Standards (ETSI in Europe, CCSA in China, etc)
  - **ETSI 103 737 Transmission requirements for *Narrowband Wireless Terminals (handset and headset)* from a QoS Perspective as Perceived by the User**
  - **ETSI 103 738 Transmission requirements for *Narrowband Handsfree Wireless Terminals* from a QoS Perspective as Perceived by the User**
  - **ETSI 103 739 Transmission requirements *for Wideband Wireless Terminals (handset and headset)* from a QoS Perspective as Perceived by the User**
  - **ETSI 103 740 Transmission requirements *for Wideband Handsfree Wireless Terminals* from a QoS Perspective as Perceived by the User**
  - **CCSA YD/T 1538 Technical Requirements and Test Methods for Acoustics Performance of Digital Mobile Terminal**

    Mobile terminals should comply to those specification when used in certain region.

Related Standards

- Industry Group (3GPP, etc)
  - **3GPP TS 51.010 (GSM 11.10) Mobile Station (MS) conformance specification; Part 1: Conformance specification**

    This specification apply only for GSM mobile speech terminals

  - **3GPP TS 26.131 Terminal acoustic characteristics for telephony; Requirements**
  - **3GPP TS 26.132 Speech and video telephony terminal; acoustic test specification**

    This specification apply only for GSM/UMTS/LTE mobile speech terminals, widely used in certification such as GCF/PTCRB/CTIA

  - **3GPP2 C.S0056 Electro-Acoustic Recommended Minimum Performance Specification for cdma2000 mobile stations**

    This specification apply only for CDMA mobile speech terminals

- Enterprise proprietary standards (AT&T, etc)
  - **AT&T audio performance standard**
  - **CMCC audio performance standard**
  - **Vodafone audio performance standard**

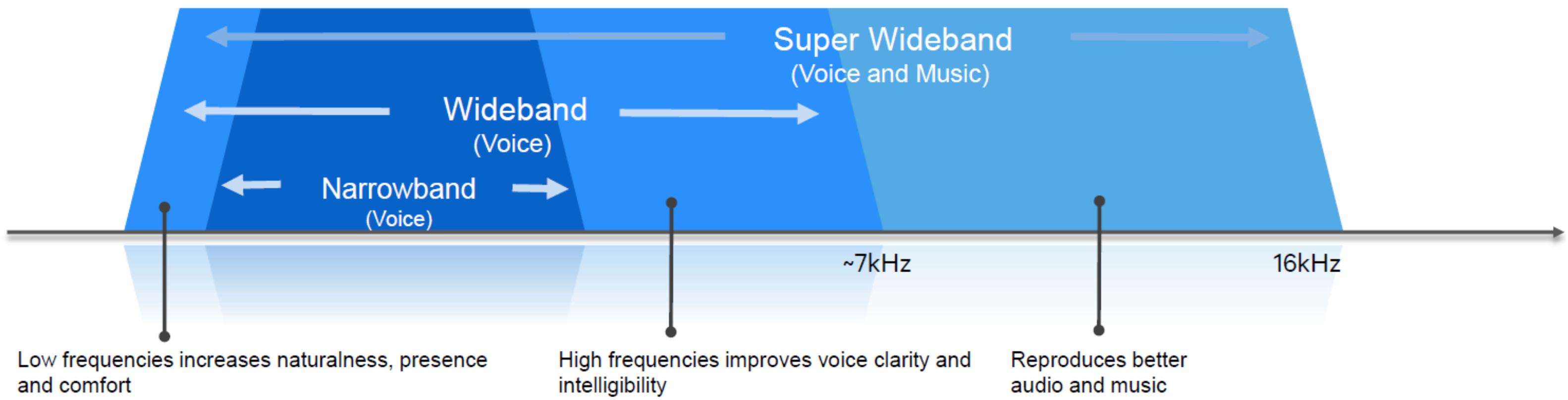
    These specification apply only for the terminals tailored for operator themselves

Outline

- The basic of electro-acoustic testing on telephone
- Test setup and test instruments
- Typical test cases of mobile phone
- Current standards in use for mobile terminals
- Future prospects

The Bandwidth extension

Terminals will support wider bandwidth of speech , brings in more superior speech quality



- **Narrowband telephony**

Transmission of a signal (either speech or data) through a telephonic network with a nominal pass-band of 300-3400 Hz.

- **Wideband telephony**

Transmission of speech with a nominal pass-band wider than 300-3400 Hz, usually understood to be 100-7000Hz.

- **Super-wideband telephony**

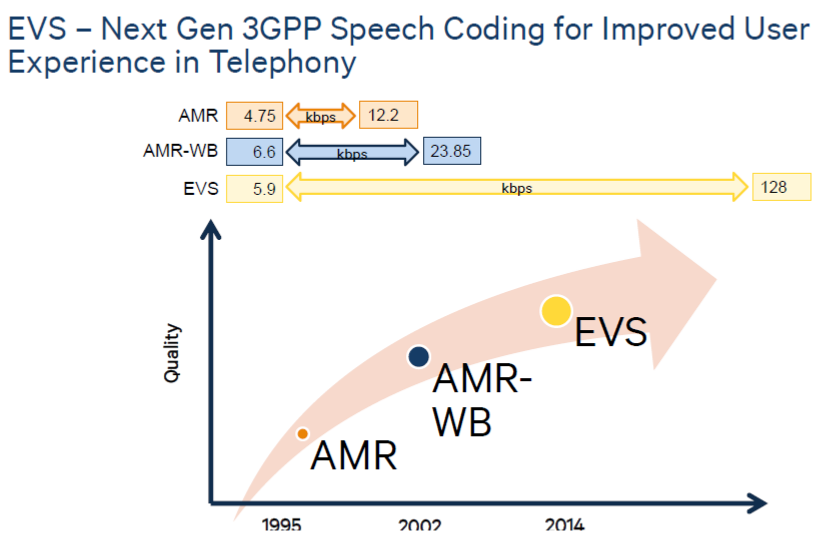
Transmission of speech with a nominal pass-band wider than 100-7000 Hz, usually understood to be 50-14000 Hz

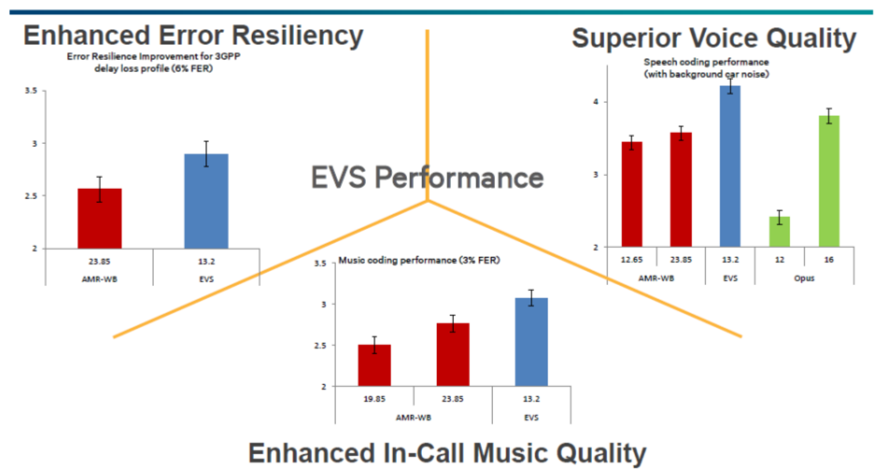
- **Fullband telephony**

Transmission of speech with a nominal pass-band wider than 50-14000 Hz, usually understood to be 20-20000 Hz. Covering all the frequency range that human can perceive.

New codec are available

Mobile terminals uses new codec to support super-wideband and fullband speech communication.





Accordingly, the test methods previously focused on narrowband and wideband speech communication will have to adapt and evolve, to obtain and reflect the actual performance under super-wideband and fullband communication modes.
It is on the way.

The end, Thank you

Hearing is a wonderful thing.

The Acoustic testing

makes *what you hear* better.

Exam questions:

1. **Is the speech quality score indicates all of the acoustic performance of mobile terminals?**
2. **What are the frequency ranges covered by narrowband and wideband speech communication?**

(Trainer information)

Trainer: Gaoxiong Yi

E-mail: yigaoxiong@caict.ac.cn

Department: CTTL-Terminals