

Translated English of Chinese Standard: GB/T45314-2025

www.ChineseStandard.net → Buy True-PDF → Auto-delivery.

Sales@ChineseStandard.net

GB

NATIONAL STANDARD OF THE
PEOPLE'S REPUBLIC OF CHINA

ICS 43.040.10

CCS T 36

GB/T 45314-2025

**Road vehicles -- Performance requirements and test
methods for hands-free communication and speech
interaction**

道路车辆 免提通话和语音交互性能要求及试验方法

Issued on: February 28, 2025

Implemented on: February 28, 2025

**Issued by: State Administration for Market Regulation;
Standardization Administration of the People's Republic of China.**

Table of Contents

Foreword	4
1 Scope	5
2 Normative references	5
3 Terms and definitions	6
4 Quality requirements for hands-free terminal calls.....	10
4.1 Time delay	10
4.2 Loudness rating	11
4.3 Sensitivity/frequency response.....	12
4.4 Unidirectional call speech quality	14
4.5 Idle channel noise.....	14
4.6 Out of band signal processing performance	15
4.7 Distortion	15
4.8 Echo suppression performance.....	16
4.9 Switch characteristics.....	19
4.10 Dual talk performance.....	20
4.11 Background noise transmission.....	22
5 Quality requirements for emergency call quality	24
5.1 Delay	24
5.2 Loudness rating	25
5.3 Sensitivity/frequency response.....	25
5.4 Idle channel noise.....	27
5.5 Echo suppression performance.....	27
5.6 Single talk activation feature for sending direction.....	28
5.7 Dual talk performance.....	28
5.8 Background noise transmission.....	29
6 Performance and functional requirements for speech interaction	29
6.1 Speech wakeup speech trigger	29
6.2 Accuracy of acoustic source localization	31
6.3 Sentence recognition accuracy	32
6.4 Interaction success rate	33
6.5 Wakeup response time.....	33
6.6 Interaction response time	34
6.7 Speech interaction terminal function.....	34
7 Test methods	34
7.1 Test conditions	34
7.2 Call quality test	41
7.3 Speech interaction performance and function test	66
Annex A (normative) Typical noise scenarios	70
A.1 Noise scenario of the hands-free terminal call quality and speech interaction	

Road vehicles -- Performance requirements and test methods for hands-free communication and speech interaction

1 Scope

This document specifies the requirements for hands-free communication and speech interaction performance of road vehicles, and describes the corresponding test methods.

This document is applicable to M₁ and N₁ vehicles equipped with car speaker hands-free communication terminals, car emergency call terminals, and car speech interaction terminals.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ITU-T G.122-1993, *Influence of national systems on stability and talker echo in international connections*

ITU-T P.51, *Simulation mouth*

ITU-T P.57, *Simulation ear s*

ITU-T P.58, *Head and torso simulator for telephony, HATS*

ITU-T P.79-2007, *Calculation of loudness ratings for telephone sets*

ITU-T P.501-2020, *Test signals for use in telephony*

ITU-T P.502-2000, *Objective test methods for speech communication systems using complex test signals*

ITU-T P.863-2018, *Perceptual objective listening quality assessment*

ITU-T P.863.1-2019, *Application guide for recommendation P.863*

ETSI EG202 396-3-2018, *Speech and multimedia Transmission Quality (STQ); Speech Quality performance in the presence of background noise; Part 3:*

[Source: ITU-T P.1100-2019, 3.2.6]

3.21 diffuse field equalization; DF

An equalization of the transmission characteristics in the receiving direction of HATS, so that the sound waves incident directly in front of HATS under diffusion field conditions have a flat frequency characteristic.

NOTE: Equalization is performed for HATS, expressed as the ratio of the average sound pressure of the undisturbed diffusion field at the HATS reference point within the 1/3 octave sub-band (i.e., the average sound pressure of the undisturbed diffusion field at the HATS reference point without placing HATS) to the average sound pressure at the ear ear-drum reference point after placing HATS.

[Source: ITU-T P.1100-2019, 3.2.4]

3.22 maximum setting of the volume

When the volume in the receiving direction is adjustable, avoid overloading the sound system and introducing non-linear relationships in the echo path, and maximize the volume of the receiving direction loudness rating.

3.23 nominal setting of the volume

To ensure that the loudness rating of the receiving direction meets the volume requirements of this document (when the receiving volume is adjustable), or the default volume of the vehicle terminal (when the receiving volume is not adjustable).

3.24 sending loudness rating; SLR

The loudness loss between the user's mouth and the electrical interface on the network end.

NOTE: The loudness loss here is defined as the weighted average of the ratio of the driving sound pressure to the measured voltage.

[Source: ITU-T P.1100-2019, 3.2.20]

3.25 receiving loudness rating; RLR

The loudness loss between the network electrical interface and the listening user's ear.

NOTE: The loudness loss here is defined as the weighted average of the ratio of the driving electromotive force to the measured sound pressure.

[Source: ITU-T P.1100-2019, 3.2.19]

3.26 junction loudness rating; JLR

It refers to the loudness loss between the network electrical interface and the short-

does not exceed 110 ms.

4.1.2 Delay in the case of using SRWT

Tested according to 7.2.1, the delay of the driver's seat shall meet the following requirements:

- When using narrowband speech, T_{rtd} shall be less than 120 ms; when using wideband speech, T_{rtd} shall be less than 130 ms;
- The T_{rtdimp} introduced by hands-free signal processing does not exceed 110 ms.

4.2 Loudness rating

4.2.1 Sending loudness rating

The sending loudness rating (SLR) of the driver's seat shall be tested according to 7.2.2.1. The SLR at the point of interconnection (POI) shall be within the range of 9 dB~17 dB.

4.2.2 Receiving loudness rating

The receiving loudness rating (RLR) of the driver's seat shall be tested according to 7.2.2.2.

If the volume in the receiving direction cannot be adjusted, the RLR at the POI shall be within the range of -2 dB~6 dB.

If the volume in the receiving direction can be adjusted, it shall meet the following requirements:

- At least one volume setting RLR meets the above requirements;
- The nominal setting of the volume RLR shall be at least 15 dB greater than the maximum setting of the volume (if the volume is adjustable) RLR;
- At least one volume setting shall ensure a signal-to-noise ratio (S/N) of at least 6 dB in any possible communication and noise environment.

4.2.3 Linearity of sending loudness rating

Tested according to 7.2.2.3, the SLR of the driver's seat shall meet the following requirements:

- When the signal level of the sending direction is relative to the nominal level by -3 dB, the deviation between the SLR and the SLR measured using the nominal value shall not exceed -0.5 dB;

4.4 Unidirectional call speech quality

4.4.1 Sending speech quality

Tested according to 7.2.4.1, the sending speech quality shall meet the following requirements:

- When using narrowband speech, the average listening opinion score (MOS_{LQO}) shall be at least 3.0;
- When using wideband speech, MOS_{LQO} shall be at least 3.6.

4.4.2 Receiving speech quality

Tested according to 7.2.4.2, the receiving speech quality shall meet the following requirements at the nominal setting of the volume:

- When using narrowband speech: MOS_{LQO} shall be at least 3.0;
- When using wideband speech, MOS_{LQO} shall be at least 3.6.

4.4.3 Stability of sending speech quality

Tested according to 7.2.4.3, the stability of the sending speech quality (ST_{MOS}) shall not be lower than 90.

4.4.4 Stability of receiving speech quality

Tested according to 7.2.4.4, the ST_{MOS} in the receiving direction shall not be lower than 90 at the nominal setting of the volume.

4.5 Idle channel noise

4.5.1 Sending idle channel noise

Test according to 7.2.5.1. For narrowband speech, the sending idle channel noise shall not exceed -64 dBm0 (P) (P weighting is specified in ITU-T O.41-1994). For wideband speech, the sending idle channel noise shall not exceed -64 dBm0 (A). There shall be no peak in the measured signal frequency domain that is more than 10 dB (including 10 dB) higher than the average level of idle channel noise spectrum. The peak value comes from the hands-free terminal in the test. If there is a peak of more than 10 dB (including 10 dB), but the idle channel noise in the direction of hands-free terminal transmission is not greater than -84 dBm0 (P) (narrowband)/-84 dBm0 (A) (wideband), then the peak can be ignored.

4.5.2 Receiving idle channel noise

Test according to 7.2.5.2. At nominal setting of the volume, the receiving idle channel noise shall not exceed -53 dBPa (A). There shall be no peak in the frequency domain of the test signal that is more than 10 dB (including 10 dB) higher than the average sound pressure of the idle channel noise spectrum. The peak value comes from the hands-free terminal in the test. If there is a peak of 10 dB or more (including 10 dB), but the receiving idle channel noise of the hands-free terminal is not greater than -73 dBPa (A), then the peak can be ignored.

4.6 Out of band signal processing performance

4.6.1 Performance of out of band signal processing for sending direction

Test according to 7.2.6.1. The out of band signal processing performance of the sending direction meets the following requirements:

- When measuring and using narrowband speech in the frequency range of 300 Hz~3400 Hz, the signal level measured at POI shall be at least 35 dB lower than the reference signal level;
- When measuring and using wideband speech in the frequency range of 200 Hz to 7000 Hz, the signal level measured at POI shall be at least 35 dB lower than the reference signal level.

4.6.2 Performance of pseudo out of band signal processing in receiving direction

Test according to 7.2.6.2. At nominal setting of the volume, the pseudo out of band signal processing performance in the receiving direction meets the following requirements:

- When measuring and using narrowband speech within the frequency range of 4600 Hz ~8000 Hz, the signal level measured at the hands-free speaker shall be at least 45 dB lower than the reference signal level;
- When measuring and using wideband speech in the frequency range of 8600 Hz~16000 Hz, the signal level measured at the hands-free speaker, both in band and out of band signals, shall not be higher than the reference signal level by more than 10 dB.

4.7 Distortion

4.7.1 Sending distortion

Test according to 7.2.7.1. The harmonic distortion in the sending direction shall not exceed 3% at 300 Hz, 500 Hz, 1000 Hz, and 2000 Hz (2000 Hz is only used for testing wideband speech).

4.8.7 Quality evaluation of speech echo

This project is optional. Test according to 7.2.8.7. At nominal setting of the volume, the score for Speech Echo Quality Evaluation (E-MOSTQO) shall not be less than 4.0.

4.9 Switch characteristics

4.9.1 Sending activation

Test according to 7.2.9.1. The minimum activated sound pressure ($L_{S, \min}$) in the sending direction shall not exceed -20 dBPa. The minimum activation establishment time ($T_{r, s, \min}$) for the sending direction shall not exceed 50 ms.

NOTE 1: $L_{S, \min}$ refers to the minimum sound pressure required to remove the insertion attenuation MRP when the transmission direction channel is in idle mode.

NOTE 2: $T_{r, s, \min}$ refers to the time required from the start of signal playback to the complete activation of the channel when the pulse signal in the sending direction reaches the minimum activation sound pressure.

4.9.2 Receiving activation

Test according to 7.2.9.2. The minimum activated sound pressure ($L_{R, \min}$) in the receiving direction shall not exceed -35.7 dBm0 (active signal part).

The minimum activation establishment time ($T_{r, R, \min}$) for the receiving direction shall not exceed 50 ms.

NOTE 1: $L_{R, \min}$ refers to the minimum sound pressure required to remove the insertion attenuation MRP when the receiving direction channel is in idle mode.

NOTE 2: $T_{r, R, \min}$ refers to the time required from the start of signal playback to the complete activation of the channel when the pulse signal in the receiving direction reaches the minimum activation sound pressure.

4.9.3 Sending attenuation range

Test according to 7.2.9.3. The attenuation range of the transmission direction ($A_{H, s}$) shall be less than 20 dB. The establishment time of the transmission direction ($T_{r, s}$) shall be less than 50 ms.

NOTE: $A_{H, s}$ refers to the difference between the minimum and maximum levels measured during the process of opening the transmission channel by adding an activation signal in the sending direction when the receiving direction is active.

4.9.4 Receiving attenuation range

Test according to 7.2.9.4. The attenuation range of the receiving direction ($A_{H, R}$) shall

dual talk transmission direction of the hands-free terminal by observing the changes in the signal level of the sending direction in the dual talk mode. Classify according to the method in Table 9.

4.10.4.2 Under the condition of nominal receiving direction volume, the dual tone level in the sending direction meets the following requirements:

- Both the sending and receiving directions use nominal signal levels. The dual talk level shall reach 2b or higher;
- The signal level in the sending direction is +6 dB relative to the nominal level. The signal level in the receiving direction is -6 dB relative to the nominal level. The dual talk level shall reach 2b or higher;
- The signal level in the sending direction is -6 dB relative to the nominal level. The signal level in the receiving direction is +6 dB relative to the nominal level. The double talk level shall reach 2b or higher.

4.10.4.3 Under the condition of maximum setting of the volume in the receiving direction (if the volume is adjustable), both the sending and receiving directions use the nominal signal level. The sending direction shall have a dual talk level of 2b or higher.

4.11 Background noise transmission

4.11.1 Background noise transmission after call establishment

Test according to 7.2.11.1. For the noise scenes A1, A2, A3, and A4 in Annex A, the significant features based on auditory related feature pattern recognition method analysis (see Annex B of ITU-T P.1100-2019) are required as follows:

- For narrowband, the excitation of the first transmission signal peak in the sending direction between 300 Hz~3400 Hz shall not exceed 15 cPa;
- For wideband, the excitation of the first transmission signal peak in the sending direction between 200 Hz and 7000 Hz shall not exceed 15 cPa;
- For narrowband and wideband, the peak value of the first transmitted signal shall not exceed 6 cPa.

4.11.2 Speech quality with background noise

Test according to 7.2.11.2. For noise scenarios A1, A2, A3, and A4 in Annex A, for speech quality with background noise, narrowband speech shall meet the requirements of Table 12, and wideband speech shall meet the requirements of Table 13.

terminal call quality and emergency call quality tests. The DUT is connected to the network system simulator through a wireless signal. The network system simulator is connected to the test system through a coaxial cable;

- b) Only use SRWT access: This connection method can be used for hands-free terminal call quality test. DUT is connected to the test system through SRWT;
- c) Using SRWT and network access: This connection method can be used for hands-free terminal call quality testing. The DUT is connected to the mobile phone via SRWT. The mobile phone is connected to the network system simulator via a wireless signal. The network system simulator is connected to the test system via a coaxial cable.

The test signal is input to the network system simulator in the form of an electrical signal or to the HATS in the form of an acoustic signal.

When using connection modes b) and c), the SRWT shall switch the corresponding narrowband/wideband coding (e.g. for Bluetooth, use CVSD coding for narrowband tests and mSBC coding for wideband tests).

When using connection method c), use the test method mentioned in Annex C to verify whether the mobile phone can be used for the test.

7.1.6 Background noise

7.1.6.1 Background noise recording

The number of recording microphones is equal to the total number of microphones in the terminal under test. The recording microphone is placed close to each microphone in the terminal under test. The recording microphone shall be installed in a way that it will not be affected by the vibration of other parts of the DUT.

7.1.6.2 Background noise playback

During the test, full-range speakers and subwoofers are arranged to reproduce the background noise in the car (see ETSI TS103 224). The schematic diagram of the speaker arrangement is shown in Figure 1. The interference of the sound transmission path between the speakers, the microphone of the terminal under test and the HATS shall be reduced. The speakers shall not resonate. The background noise playback system is time-synchronized with the test system.

The loudspeaker group shall be equalized so that the reproduced sound field of each recording microphone is equal to the recorded noise sound field:

- The level difference shall be within ± 1 dB;
- In the frequency range of 100 Hz~1000 Hz, the amplitude of the complex coherence (normalized cross-correlation spectrum) shall be greater than 0.9

(measured in 1/3 octave);

- The phase of the complex coherence shall be accurate to within $\pm 10^\circ$ in the frequency range of 100 Hz~1000 Hz. In the frequency range of 1000 Hz~1500 Hz, the accuracy shall be within $\pm 30^\circ$ (measured in 1/3 octave band);
- The difference between the original reference noise amplitude spectrum and the simulated noise amplitude spectrum of each recording microphone shall be within ± 3 dB in the frequency range of 50 Hz~10000 Hz, and within ± 6 dB in the frequency range of 10000 Hz~16000 Hz (measured in 1/3 octave band);
- In the frequency range of 50 Hz~20000 Hz, the average spectrum accuracy of all recording microphones shall be within ± 3 dB.

7.2 Call quality test

7.2.1 Delay

7.2.1.1 Delay in the sending direction

Place the HATS in the driver's seat for testing. Equalize the output of the simulation mouth. The sound pressure level at the HFRP is -25.7 dBPa. The test signal uses the CSS of ITU-T P.501-2020 as the source signal. Compare the test signal at the POI output interface with the source signal. Use the cross-correlation function method to determine the total transmission delay in the sending direction (the time difference at the maximum point of the cross-correlation function is determined as the delay). The delay is in ms. After removing the test system delay (T_{system}), the delay in the sending direction (T_s) is obtained.

NOTE 1: The pseudo-random noise portion of the CSS signal (PN sequence) occupies a duration greater than the total transmission delay in the transmit direction, using a PN sequence of 16000 samples (with a sampling rate of 48000 Hz) or other equivalent sequences.

NOTE 2: T_{system} is a known parameter that depends on the signal transmission method and the network system simulator.

7.2.1.2 Delay in receiving direction

Place the HATS at the driver's seat for testing. Perform DF or FF on the simulation ear. The level of the test signal at the POI is -16 dBm0. The test signal uses the CSS of ITU-T P.501-2020 as the source signal. The volume in the DUT receiving direction is set to the nominal setting of the volume. Compare the test signal at the ear-drum reference point (DRP) with the source signal. Use the cross-correlation function method to determine the total transmission delay in the receiving direction (the time difference at the maximum point of the cross-correlation function is determined as the delay). The delay is in ms. After removing T_{system} , the delay in the receiving direction (T_R) is obtained.

transmission direction sensitivity.

- For narrowband speech calls, the sending sensitivity is calculated for each of the 14 frequencies (bands 4~17) in Table 1 of ITU-T P.79-2007. For wideband speech calls, the sending sensitivity is calculated for each of the 20 frequencies (bands 1~20) in Table A.2 of ITU-T P.79-2007. When calculating, the average value of the electrical signal at the POI output interface and the reference signal at the MRP in each sub-band is compared.
- Sensitivity is expressed in dBV/Pa. For narrowband, according to formula 5-1 in ITU-T P.79-2007, frequency bands 4~17, $m=0.175$, calculate the SLR based on the weight of the sending direction in Table 1. For wideband, according to Annex A of ITU-T P.79-2007, frequency bands 1~20, calculate the SLR.

7.2.2.2 Receiving loudness rating

During the hands-free terminal call quality test, place the HATS in the driver's seat for testing. During the emergency call quality test, HATS will be placed in the driver's seat, copilot's seat, and two second row outer passenger seats for testing. Simulate ear for DF or FF. The test signal uses the Chinese speech signal single talk sequence in ITU-T P.501-2020, with a level of -16 dBm0 at POI (signal amplitude refers to the average value of the entire signal duration). The test method is as follows.

- a) Set the volume in the receiving direction of the tested terminal to the nominal setting of the volume.
- b) The test uses balanced output signals from two simulation ears. The balanced output signal of each simulation ear is the average power over the entire analysis time. The left and right ear signals are the sum of the voltages of each 1/3 octave band.
- c) For narrowband speech calls, calculate the receiving sensitivity for each of the 14 frequencies (bands 4~17) in Table 1 of ITU-T P.79-2007. Wideband speech communication calculates the receiving sensitivity for each of the 20 frequencies (bands 1~20) in Table A.2 of ITU-T P.79-2007. When calculating, the average signal level of each frequency band refers to the signal level of the reference signal measured in each frequency band.
- d) Sensitivity is expressed in dBPa/V. RLR shall be calculated according to Annex A of ITU-T P.79-2007 (without LE factor).
- e) Subtract the correction value of 8 dB from the measured value to obtain the final result.
- f) Repeat steps b) ~ d) at the maximum setting of the volume of the tested terminal.

7.2.2.3 Linearity of sending loudness rating

This is an excerpt of the PDF (Some pages are marked off intentionally)

Full-copy PDF can be purchased from 1 of 2 websites:

1. <https://www.ChineseStandard.us>

- SEARCH the standard ID, such as GB 4943.1-2022.
- Select your country (currency), for example: USA (USD); Germany (Euro).
- Full-copy of PDF (text-editable, true-PDF) can be downloaded in 9 seconds.
- Tax invoice can be downloaded in 9 seconds.
- Receiving emails in 9 seconds (with download links).

2. <https://www.ChineseStandard.net>

- SEARCH the standard ID, such as GB 4943.1-2022.
- Add to cart. Only accept USD (other currencies - <https://www.ChineseStandard.us>).
- Full-copy of PDF (text-editable, true-PDF) can be downloaded in 9 seconds.
- Receiving emails in 9 seconds (with PDFs attached, invoice and download links).

Translated by: Field Test Asia Pte. Ltd. (Incorporated & taxed in Singapore. Tax ID: 201302277C)

About Us (Goodwill, Policies, Fair Trading...): <https://www.chinesestandard.net/AboutUs.aspx>

Contact: Wayne Zheng, Sales@ChineseStandard.net

Linkin: <https://www.linkedin.com/in/waynezhengwenrui/>

----- The End -----