ETSI
3rd Speech Quality Test Event

“Gateway Pie”

Conversational Quality Overview

Annex

to the individual Test Report AudioCodes Ltd.

ETSI Plugtests™ Service
HEAD acoustics GmbH

Conversational Quality Overview - Introduction

The tests carried out during the 3rd ETSI Speech Quality Test Event cover all conversational speech quality aspects. Listening speech quality, echo measurements, double talk performance and background noise transmission are considered. A variety of different test signals and speech samples were used. The tests were carried out applying different signal levels and various kinds of analysis methods. Limitations are of course given by the choice of test conditions like the network impairments or echo path realizations. The tests and the test setup were discussed and chosen in order to reproduce a wide range of defined and reproducible conditions.

In order to provide a condensed overview about the results for each gateway under tests for all speech quality aspects a graphical result representation is derived. The focus of this conversational speech quality representation is

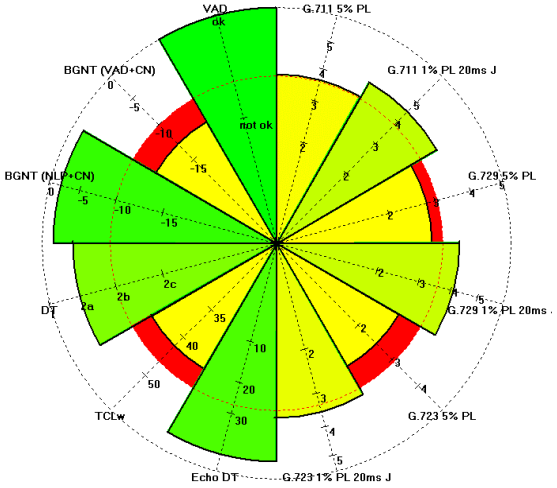
- To provide a condensed, “quick and easy to read” overview about the current implementation.
- To provide a variety of measurement results and compare them to the recommended values and numbers in current ITU-T or ETSI standards.
- To provide a comparison to average results from all manufacturers participating in this event.
- To give an indication about strength and weakness of the different implementations.
- To provide detailed enough information for engineering and development in order to improve the performance.

The results are summarized in one diagram, best described as a “Quality Pie”. Each pie slice represents a transmission performance parameter like the codec performance under 5% packet loss, the echo attenuation under single talk conditions, the quality of background noise transmission or others. The size of each slice represents a measure for the quality of this parameter. Bigger slices indicate a better performance.

All relevant analyses for this representation are given in this report above. The following example of this “Gateway Pie” does not represent an existing gateway implementation, it is used as an example in order to introduce this result overview.

The following assumptions were made:

- Each parameter is represented by a pie slice.
- The pie slices are independent from each other. Interaction between different parameters like the echo perception due to the combination of echo attenuation and speech distortions (introduced by speech coders) are not considered.
- The size of each slice directly correlates to quality. The pie slice size is area equivalent.

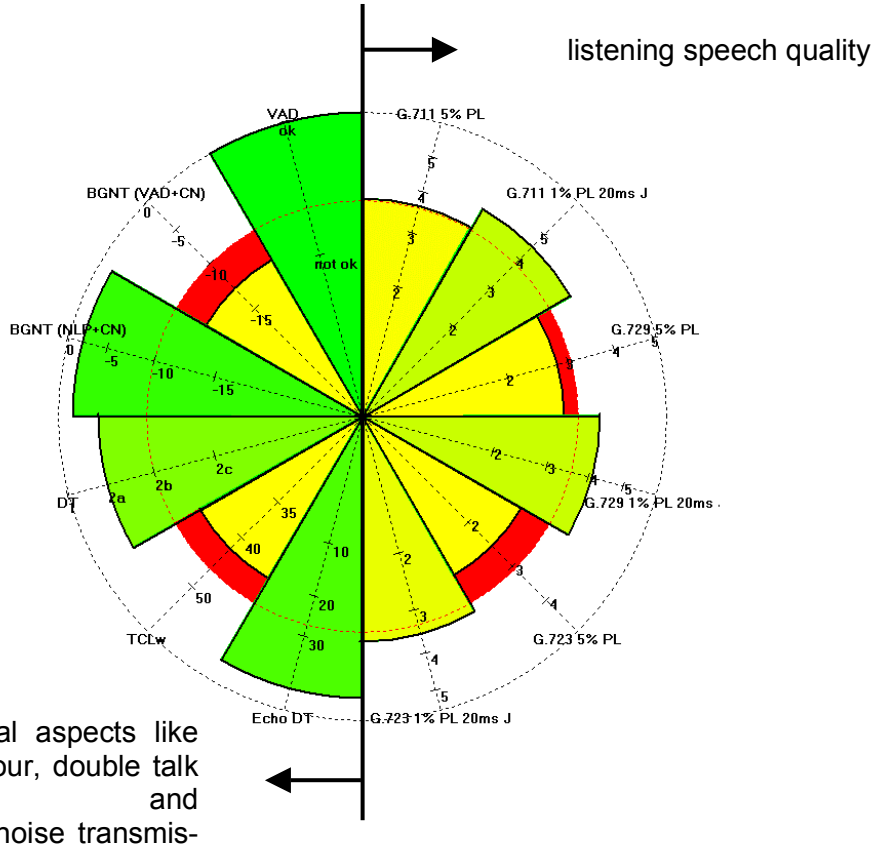


Example “Gateway Pie”

- The minimum requirement for a parameter or the average results from all manufacturers participating during the event is indicated by an inner red circle. If the measured parameter exceeds the recommended requirement or indicates a quality better than the average performance during the test event, the red circle is not visible and overlapped by the pie slice.
- In addition the size of a pie slice is colour coded from yellow (low quality scores or low values) to green (high quality scores or high values).
- The axis scale of each pie slice is parameter dependent.

The following example introduces this conversational speech quality representation and explains, "how to read it". The example does not representing an existing gateway.

Example:



conversational aspects like echo behaviour, double talk performance and background noise transmis-

The right hand half of the pie represents the listening speech quality for the different speech codecs. These results consider the influence of packet loss and jitter (G.711, G.729 and G.723 under test condition 5a and 7a).

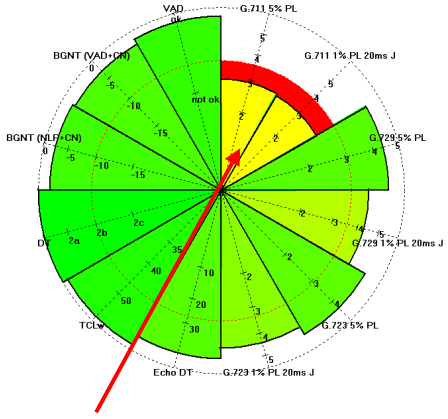
The left hand side represents the conversational aspects like

- Echo performance during double talk and single talk ("Echo DT", "TCL_w").
- The double talk performance ("DT"), characterization in accordance to ITU-T Recommendation P.340 [25] (note that these references can be found in the *Anonymized Test Report – "Gateways"*).
- The quality of background noise transmission ("BGNT(NLP+CN)"), modulation introduced by the echo suppression unit and its associated comfort noise injection during the application of far end signal.

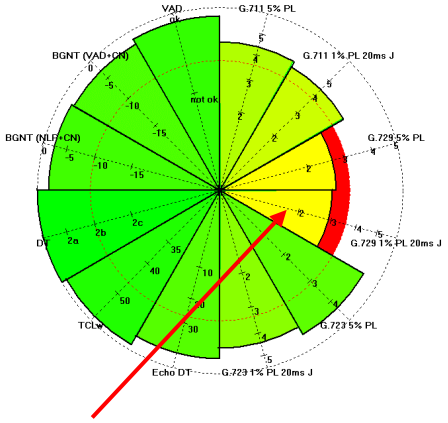
- The quality of background noise transmission in one way scenarios ("BGNT(VAD+CN)"), modulation caused by voice activity detection of comfort noise injection under single talk conditions.
- The performance of the implemented VAD respectively automatic gain control ("VAD").

The following examples explain each transmission quality parameter ("pie slice") with its scaling and requirement in detail. Again these examples are not derived from real existing gateways.

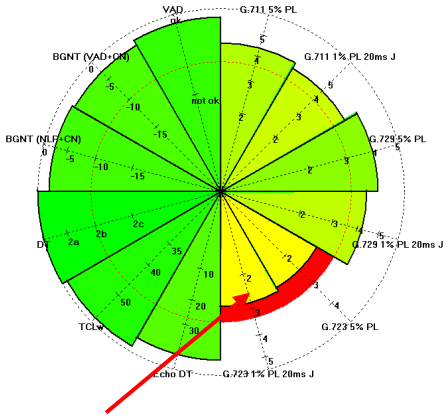
Examples (listening speech quality):



"G.711 listening speech quality below average results under both condition 5a and 7a"



"G.729 listening speech quality below average results under both condition 5a and 7a"



"G.723 listening speech quality below average results under both condition 5a and 7a"

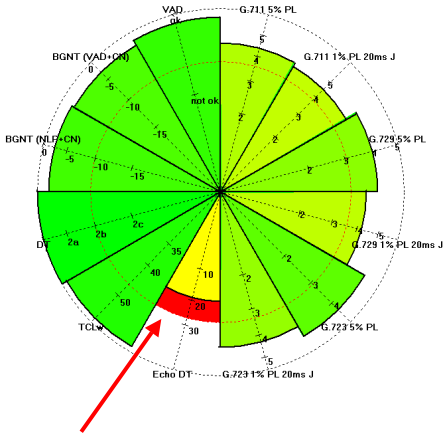
The listening speech quality result for each speech coder is represented by two slices, one for the packet loss condition 5a (5%), one for the jitter condition 7a (20ms jitter, 1% packet loss). The values are taken from the MOS-LQO result tables in the individual reports for the G.711, G.729 and G.723 speech coder. Each axis is scaled between 1 and 5 representing the MOS scale.

The limit (radius of the red circle) is given by the average MOS-LQO result over all participants under this test condition (average result from **table 5.2**, **5.4** and **5.6,** *Anonymized Test Report – "Gateways"*). It should be considered that this limit is codec dependent, thus the limits are different for the three speech coders.

Examples (echo during double talk and single talk):

The echo attenuation during double talk was measured using the AM/FM modulated test signal. These tests were carried out with a 40 dB ERL and 6 dB ERL echo path.

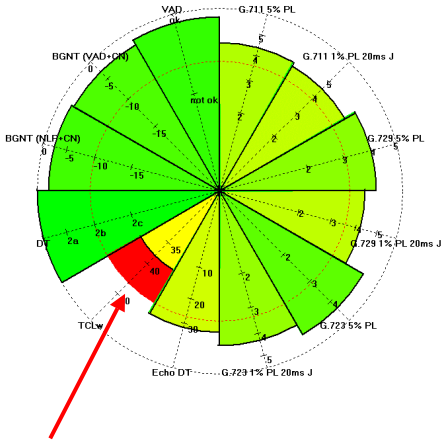
The minimum attenuation (indicated by the inner red circle) is 27 dB. This value, derived from subjective tests can be found in ITU-T P.340 [25]. 27 dB echo attenuation during double talk would lead to a full duplex characterization assuming a 100 ms one-way delay in the network. This value can be regarded as a minimum requirement.



"Echo attenuation under double talk conditions lower than recommended"

The echo loss results can be found in **table 5.8** in the *Anonymized Test Report – "Gateways"*. The relevant results for this representation are taken from the 6 dB and the 40 dB ERL measurement.

The lower value from both measurements is used for the pie. The requirement represented by the inner red circle is 46 dB.



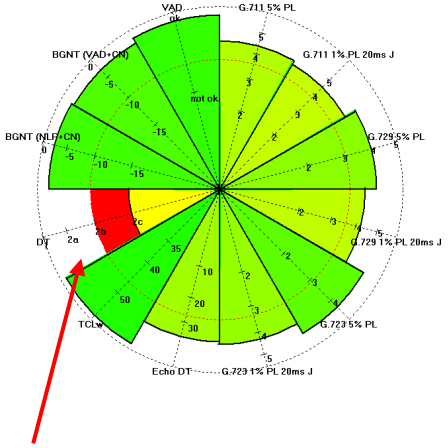
"Echo attenuation according to G.122 under single talk condition below 46 dB"

Example (attenuation during double talk, characterization):

The double talk performance is influenced by the attenuation inserted during a double talk period. Many double talk tests were carried out during the event with a 40 dB ERL and 6 dB ERL echo path.

The test results for the pie slice can be found in **figure 5.27** and **5.32 (***Anonymized Test Report – "Gateways"*). The analyses in these figures represent extreme conditions (high level differences) of the whole sequence.

In accordance to listening examples recorded during the event using real speech, the analyzed sequence from the test signal is chosen between 5.55 s and 8.35 s, where the receive signal and the near end signal are applied with the same level.



"Double talk performance influenced by level variation"

The listening recordings were also carried out under these conditions (same level in both directions).

The level of the transmitted signal is referred to the near end signal level (double talk signal) and analyzed vs. time. The average level difference is used to classify the double talk performance.

These results correlate to the listening examples recorded during the event. Level differences between 0 and 3 dB lead to a type 1 characterization (full duplex capability). This was achieved by all implementation under test.

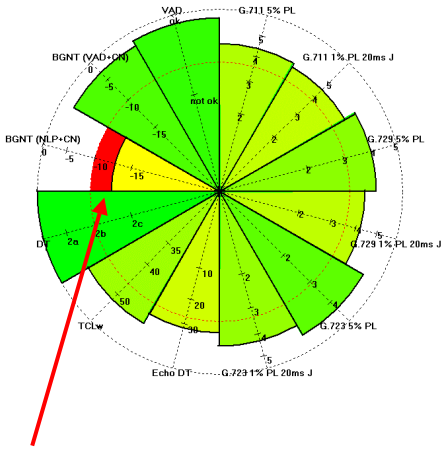
More detailed double talk performance tests were carried out during the event with higher test signal level variations (see **figure 5.27**, **5.28**, **5.32** and **5.33,** *Anonymized Test Report – "Gateways"*). These analysis results may provide useful information for manufacturers in order to optimize double talk performance.

Example (quality of background noise transmission with far end signal):

During the application of far end signals the echo suppression unit may introduce audible and disturbing noise modulation (level variation).

The relevant tests can be found in **figure 5.34** and **5.24** in the *Anonymized Test Report – "Gateways"*. The level difference between the transmitted signal with and without the application of far end signals is measured.

This difference should not exceed 10 dB, neither for the pub noise nor for the café noise.

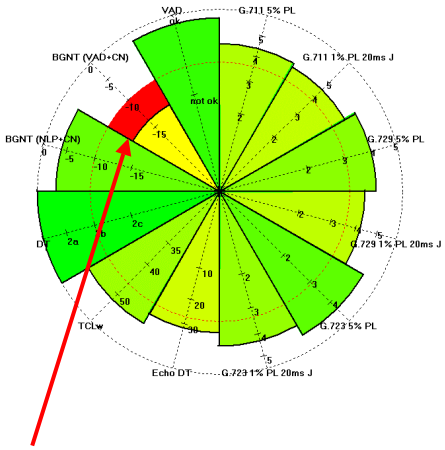


"Background noise modulation introduced by echo suppression and/or comfort noise generation too high"

Example (quality of realistic background noise transmission):

Realistic background noise scenarios like the pub noise or the café noise used during this test event should be transmitted without significant level variation. The relevant tests can be found in **figure 5.21** and **5.22** in the *Anonymized Test Report – "Gateways"*. The level difference between the transmitted signal with and without VAD is measured.

This difference should not exceed 10 dB, neither for the pub noise nor for the café noise.

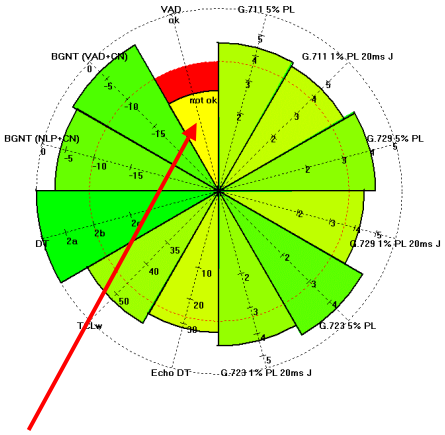


"Background noise modulation introduced by VAD or comfort noise generation too high"

Example (VAD and AGC test):

The level of a transmitted test signal should follow the original test signal level, if VAD is enabled. Comfort noise –if implemented- should be level adaptive. The relevant analyses can be found in **figure 5.14** in the *Anonymized Test Report – "Gateways"*.

The level of the transmitted signal should meet the tolerance scheme in **figure 5.14**. This tolerance scheme was derived from test results during the 2nd ETSI SQTE [21].

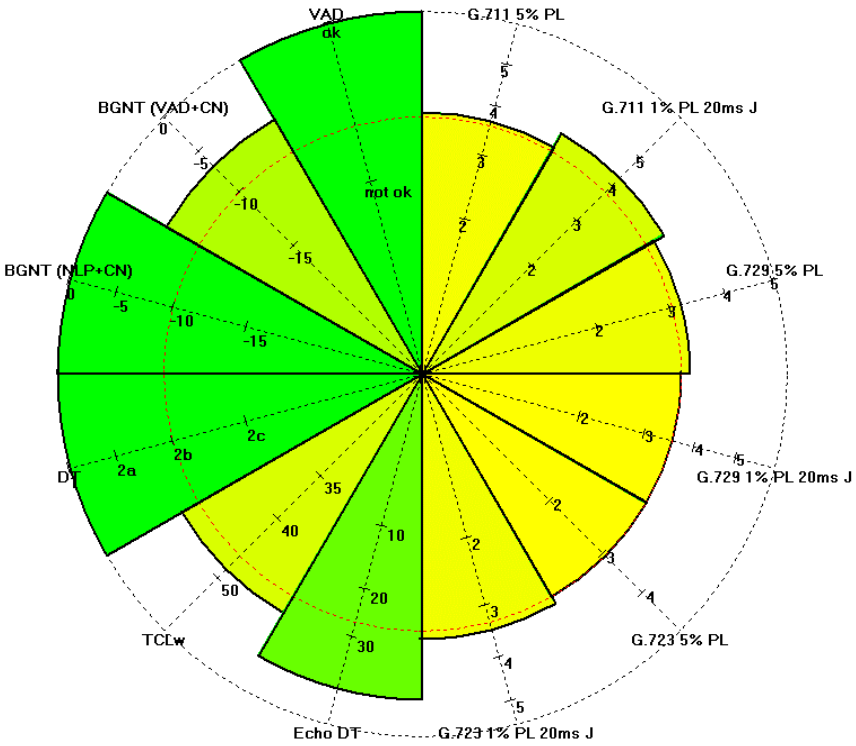


"Level of transmitted signal violates the tolerance scheme"

Summary – “Gateway Pie”

The speech quality test results for the Audiocodes gateway **Mediant 2000** are summarized and condensed to this graphical representation (“Quality Pie”). It covers all conversational aspects tested during the event like listening speech quality, echo measurements, double talk performance and quality of background noise transmission.

This characterization is based on the tests carried out during the 3rd Speech Quality Test Event. Due to the nature of a test event, the tests can not lay claim to completeness.

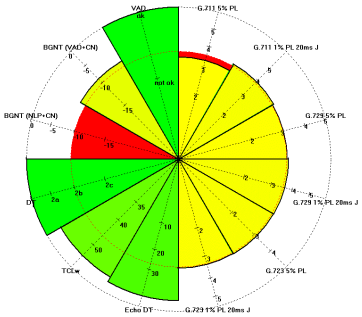


Characterization:

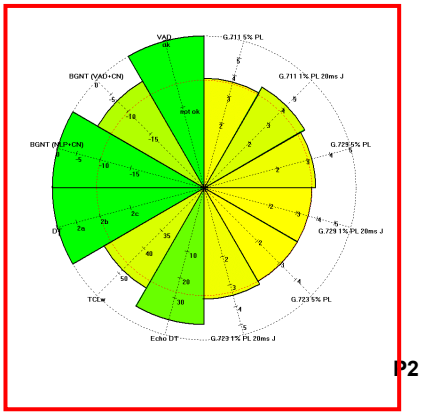
- The listening speech quality scores are comparable or slightly higher (G.711, jitter) than the average scores.
- The echo attenuation under single and double talk conditions exceeds the recommended value. Double talk performance is characterized as “full duplex” for comparable near end and far end signal levels.
- The activation of echo suppression does not lead to disturbing noise modulation (tested with infinite ERL).
- VAD and comfort noise generation do not significantly modulate the transmitted pub and café noises.

This figure, taken from the ***Anonymized Test Report "Gateways"***, provides an overview over all "Gateway Pies". Note that the gateways are analyzed in randomized order.

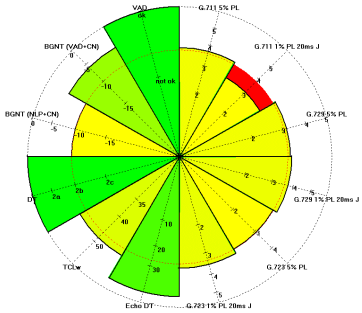
The AudioCodes gateway is represented by P2.



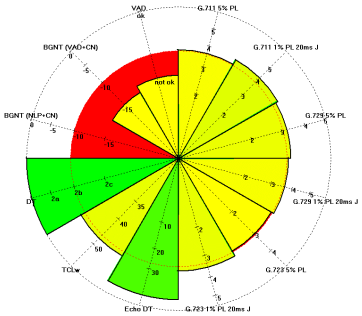
P1



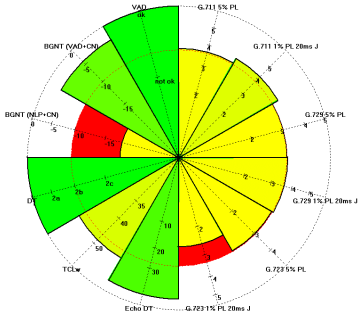
P2



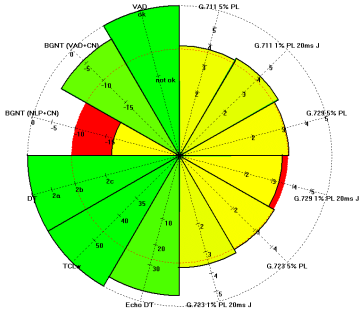
P3



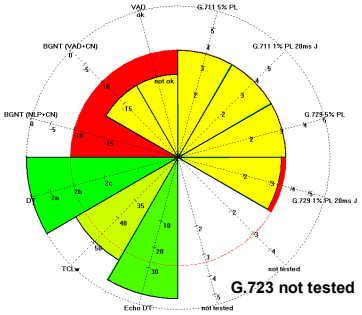
P4



P5

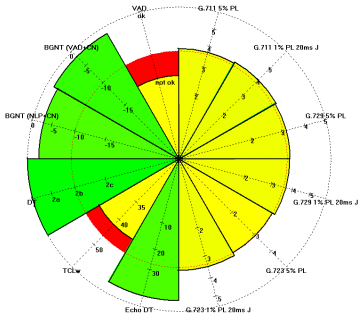


P6

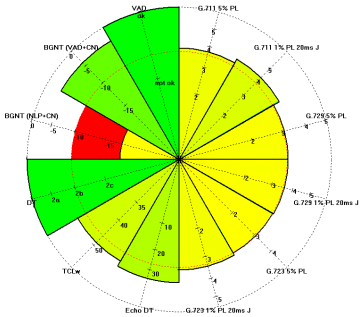


G.723 not tested

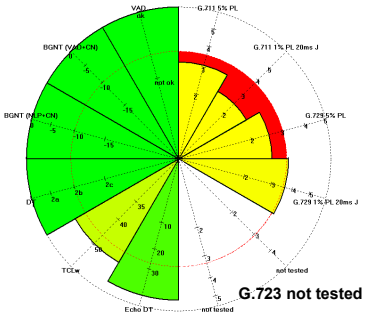
P7



P8



P9



G.723 not tested

P10