

Appendix A

Short Conversation Test (SCT)

Scenario: British Rail Travel Information

Your name: *Thompson*



Intended journey: London King's Cross → Inverness



Date: July 20th

Night train connection, sleeper

Through train if possible



Departure time : _____ h

Arrival time : _____ h

Train number : _____



Reservation : One berth in a sleeping-car

First class

Non-smoker



What time does the buffet car start serving breakfast?

Scenario 1: British Rail Travel Information

Your name: *British Rail travel service*



Timetable: London ⇒ Inverness

<i>Train number</i>	IC 1127	IC 829	IC 1511
London King's Cross	<i>Dep.</i> 19:18 h	20:50 h	22:09 h
York	<i>Arr.</i>	21:49 h	
<i>Train number</i>		EN 225	
York	<i>Dep.</i>	22:44 h	
Inverness	<i>Arr.</i> 4:39 h	6:30 h	7:16 h
	Sleeper only 2nd class	Sleeper only 2nd class	Sleeper 1st and 2nd class



Reservation: Name : _____

Number of berths : _____

Class : _____

Smoker

Non-smoker





Appendix B

Random Number Verification Task (RNVT)

Example scenarios for random number verification tasks (RNVT)

Test subject 1

Instructions: “Your conversation partner is also provided with such a list. Some of the numbers in your list do not correspond with those of your conversation partner. Find the wrong numbers as quickly as possible by taking turns reading them line by line. Acknowledge by saying “yes” or “no”, and cross out the wrong numbers. You will read the bold red numbers and your conversation partner will read the non-bold blue ones”.



You are the "caller".

18	88	80	74	55	7
15	29	14	37	17	82
20	95	36	77	34	83
46	84	30	67	25	99
28	27	36	96	60	97
55	10	87	53	43	98

Example scenarios for random number verification tasks (RNVT)**Test subject 2**

Instructions: “Your conversation partner is also provided with such a list. Some of the numbers in your list do not correspond with those of your conversation partner. Find the wrong numbers as quickly as possible by taking turns reading them line by line. Acknowledge by saying “yes” or “no”, and cross out the wrong numbers. You will read the bold blue numbers and your conversation partner will read the non-bold red ones”.



Please wait to be called.

18	84	80	74	55	7
15	29	14	67	17	82
36	95	36	77	53	83
46	88	30	37	25	99
28	27	20	96	60	97
55	10	87	34	43	98

Appendix C

Test Instructions

Test instructions—translated from german

Analysis of speech quality in a Conversational situation

Thank you very much for taking part in this experiment! Please take the time to read the instructions thoroughly. Should you have any questions please do not hesitate to ask the experiment supervisor.

You are taking part in a conversational experiment, in which the properties and the perception of a conversation are to be evaluated. To do this you and you conversational partner will be put into a conversational situation. In this situation different conditions will be presented in three phases. The phases are the following: (1.) a short conversational scenario, (2.) speaking and listening and (3.) the interaction. After every phase you are asked to evaluate your perception and the properties of the phase. The exact procedure of the individual phases will be described in detail later, and will be made clear in a test run.

The characteristics of each condition are to be evaluated with the help of different scales. Please familiarize yourself with the scales and how to use them now:

Each scale has a term at each end (e.g. not noisy/noisy). You are to evaluate to what extent the characteristics of a condition can be described by the terms on the scale. There are two different kinds of scales. The scale for the overall quality has describing terms for every point while the scales for the characteristics (e.g. not noisy/noisy) only have terms at their endpoints. The usage of these scales is analogous and is explained in detail for the overall quality.

The **overall quality** of a condition is to be evaluated. If you are of the opinion the condition is overall bad, put down your cross at the following position:



If you think the overall quality of the condition is excellent, put down your cross at the position “excellent”:



For your evaluation you can use the whole scale freely. The markings on the scale are meant to provide points of reverence for you and to support you in your evaluation. You can even use the spaces in between the markings if you do not wish to settle for one of the markings.

You can always use the “overflow areas” beyond the terms if you feel the terms are not sufficient for your evaluation:



The scales are used in an analogue way for the different characteristics. Overall there are seven characteristics to evaluate.

1. Noisiness

The first scale has the terms “not noisy” and “noisy” and looks the following:



With this scale you are to evaluate the **noisiness** of what you heard in a condition. The terms “not noisy” and “noisy” can be described with the terms *noisless* and *not hissing*, and *noisy* and *hissing* respectively.

2. Discontinuity

The **discontinuity** of the heard sound in a condition is to be evaluated with the second scale:



The term “continuous” means that the heard sound in a condition is completely *even*, *firm*, *not chopped* and *not frayed*. “Discontinuous” can be describes with terms like *uneven*, *wobbly*, *chopped* or *frayed*.

3. Coloration

The **coloration** of the heard sound in a condition is described by the third scale:



“Uncolored” means that the heard sound in a condition is *direct*, *close*, *full* and *not nasally*. “Colored” means the heard sound is *indirect*, *far*, *thin* and *nasally*.

4. Loudness

The fourth scale describes the **Loudness** of the heard sound in a condition:



If the heard sound is neither too loud nor too quiet then the volume is optimal. If that is not the case then the volume is not optimal.

5. Impact of one’s own voice

The **impact on your speaking through the hearing of your own voice** is to be evaluated with the fifth scale:



In some of the conditions you will be confronted with back coupling of your own voice. This means that you will hear your own voice. With this characteristic you are to evaluate if this does or doesn’t negatively impact your speaking. More to the point you are to evaluate if speaking while hearing yourself is *not distracting, not irritating, fluid* and *doesn’t take concentration* or if it is *distracting, irritating, not fluid* and *takes concentration*.

6. Degradation of one’s own voice

The sixth scale is for the evaluation of the **degradation of the own voice** while having back coupling:



“Own voice not degraded” means that you hear your own voice unaltered e.g. you hear your voice *not distorted, echofree* and with *optimal volume*. “Own voice degraded” can be described with terms like *distorted, echoy, thin* or *not optimal volume*

7. Interactivity

The seventh scale is used to evaluate the **interactivity** in a condition:



“Easy to interact” means that the interaction between you and your conversational partner is easy. That means you have an *effective, pleasant, easy, and agile* interaction. “Hard to interaction”, on the other hand, means that the interaction is *ineffective, unpleasant, hard and sluggish*.

For each scale applies: Use as described for the scale for “overall quality” (see above). Now that you have familiarized yourself with the meaning of the characteristics, the phases and the procedure of the test will be described.

Phase 1—Conversation

Explanation:

In this phase you and your partner simulate a telephone conversation. One of you will be the caller, who wants something and calls a company/organization/institution. The other one will be the company/organization/institution.

Explanation of symbols:

As the caller you have the following symbols:



This symbol means: You are the caller
Please wait until the experiment supervisor asks you to start the first conversation



Next to this symbol is the reason for your call
e.g.: I want to buy a ticket!



Next to this symbol are the conditions that should be incorporated into the exchange of information
e.g.: I want to buy a ticket! → BUT if possible cheap!



Next to this symbol you are to note down all information that you need from your telephone partner



Next to this symbol are all the information that your partner needs and that you should give him at some point during the conversation



Next to this symbol is a question to which neither you nor your conversational partner have any information. You are to discuss this question briefly and come to a satisfactory conclusion together

If you get called you have to following symbols:



This symbol means: You are getting called. Wait until you hear the ringtone and then pick up



Next to this symbol is information from which you are to sort out the information that your partner needs to have
e.g.: Prices for train tickets for adults, students, children, seniors, etc.



Next to this symbol you are to record all information that you need from your conversational partner

Procedure:

Read the Information that you are given once before starting the conversation. Are you the caller, call your partner. Play through the scenario and hang up the conversation. Now click “Next”. First you evaluate the overall quality of the conversation and then the conversation regarding the seven characteristics.

Are you the one who gets called, please wait until you hear the ringtone. When you do, pick up. Play through the scenario with your partner and wait until they hang up the conversation.

Click on “Next”. First you evaluate the overall quality of the conversation and then the conversation regarding the seven characteristics.

Phase 2—Listening and Speaking:

Explanation:

In this phase either you read two sentences to your partner or your partner reads two sentences to you. Afterwards you swap around.

Procedure:

Should you have two sentences before you: Call your partner and read both sentences. Hang up the conversation. Click “Next”. Now you need to evaluate the overall quality of your speaking experience. Afterwards you need to evaluate your speaking pertaining the two characteristics “Impact of the own voice” and “degradation of the own voice”.

Should you have no sentences to read: Wait until you are called and pick up the conversation. Listen to the two sentences and wait until your partner hangs up. Click “Next”. Evaluate the “overall quality” of your listening experience first, followed by the four conditions “noisiness”, “coloration”, “continuity” and “volume”.

Phase 3—Interaction:

Explanation:

In this phase you and your partner are to do a number verification.

Procedure:

You will see four series of numbers before you. Two are **bold** and two are not. You are to read the bold number sequences to your partner. Should the first of the

four sequences on your screen be bold, please call your partner. Do the number verification in an alternating fashion and when all four sequences have been read hang up the conversation. Wait after every number in the sequence (e.g. 24) for a confirmation of your partner in form of a “yes” when they have the same number or a “no” if they should have a different number. After the conversation, click “Next”. Please evaluate the “overall quality” of the interaction and afterwards evaluate the characteristic “interaction”.

Overall procedure:

Summarized the test goes for every condition the following way (eleven conditions in total):

call → phase 1 → hang up → Next → overall quality → evaluation 7 characteristics → call → phase 2 (speaking/listening) → hang up → Next → overall quality → 2/4 characteristics → call → phase 2 (speaking /listening) → hang up → Next → overall quality → 2/4 characteristics → call → phase 3 → hang up → Next → overall quality → interaction.

With a few of the conditions you could have the feeling that you have evaluated them before. This is not the case. Please evaluate every condition independently from all your other evaluations. Try not to remember how you evaluated other “similar” evaluations before, but evaluate every characteristic of every condition individually.

Please evaluate the scales quickly and intuitively. This experiment is purely subjective in its nature. There are no right or wrong answers. **Not you but the system is being tested.** Only your personal opinion matters for this experiment.

If you have any more questions please don’t hesitate to ask the experiment supervisor. Have fun! ☺

References

1. Köster F, Möller S, Antons J-N, Arndt S, Guse D, Weiss B (2014) Methods for assessing the quality of transmitted speech and of speech communication services. *Acoust Aust* 42(3):179–184
2. Wältermann M (2012) Dimension-based quality modeling of transmitted speech. Springer, Berlin
3. Côté N (2011) Integral and diagnostic intrusive prediction of speech quality. Springer, Berlin
4. Lyons J (1977) *Semantics*, vol 2. Cambridge University Press, Cambridge. Cambridge Books Online
5. Denes P, Pinson E (1993) *The Speech Chain*. Ser. Anchor books. Worth Publishers, New York
6. Möller S (2000) *Assessment and prediction of speech quality in telecommunications*. Kluwer, Boston
7. Richards D (1973) *Telecommunication by speech: the transmission performance of telephone networks*. Butterworths, London
8. Vary P, Heute U, Hess W (1998) *Digitale sprachsignalverarbeitung*. Teubner, Stuttgart
9. Peterson GE, Barney HL (1952) Control methods used in a study of vowels. *J Acoust Soc Am* 24(2):175–184
10. O’Shaughnessy D (2000) *Speech communications: human and machine*. Institute of Electrical and Electronics Engineers, New York
11. Association IP (1999) *Handbook of the international phonetic association: a guide to the use of the international phonetic alphabet*. ser. A Regents publication. Cambridge University Press, Cambridge
12. Hess W (1976) A pitch-synchronous digital feature extraction system for phonemic recognition of speech. *IEEE Trans Acoust Speech Sig Process* 24:14–25
13. Deng L, O’Shaughnessy D (2003) *Speech processing: a dynamic and optimization-oriented approach*. Ser. Signal Processing and Communications. Taylor & Francis, London
14. Jekosch U (2005) *Voice and speech quality perception: assessment and evaluation*. Springer Science & Business Media, Berlin
15. Berger J (1998) *Instrumentelle Verfahren zur Sprachqualitätsschätzung: Modelle auditiver Tests*. Shaker Verlag, Kiel
16. ITU-T Recommendation I.430 (1995) *Basic user-network interface—Layer 1 specification*. International Telecommunication Union, Geneva
17. Lewcio B (2013) *Management of speech and video telephony quality in heterogeneous wireless networks*. Springer Publishing Company, Incorporated, Berlin

18. 3GPP TS 45.-series 3rd generation partnership project. <http://www.3gpp.org/DynaReport/45-series.htm>
19. 3GPP TS 25.-series 3rd generation partnership project. <http://www.3gpp.org/DynaReport/25-series.htm>
20. 3GPP TS 36.-series 3rd generation partnership project. <http://www.3gpp.org/dynareport/36-series.htm>
21. Raake A (2006) *Speech quality of VoIP assessment and prediction*. Wiley, Chichester
22. Zwicker E (1982) *Psychoakustik*. Springer, Berlin
23. Blauert J, Xiang N (2009) *Acoustics for engineers*. Springer, Berlin
24. Fastl H, Zwicker E (2007) *Psychoacoustics: facts and models*, 3rd edn. Springer, Berlin
25. Hauenstein M (1997) *Psychoakustisch motivierte Maße zur instrumentellen Sprachgütebeurteilung*. Shaker Verlag, Kiel
26. Allen J (1994) How do humans process and recognize speech? *IEEE Trans Speech Audio Process* 2(4):567–577
27. Ogden CK, Richards IA (1989) *Meaning Of meaning*. Mariner Books, New York
28. Nöth W (2000) *Handbuch der Semiotik*, 2nd edn. 0.4emJ.B. Metzlersche Verlagsbuchhandlung
29. Baddeley A (1997) *Human memory: theory and practice*. Psychology Press, Exeter
30. Demany L, Semal C The role of memory in auditory perception. In: *Auditory perception of sound sources*. Springer Science and Business Media, pp 77–113
31. Cowan N (1984) On short and long auditory stores. *Psychol Bull* 96(2):341–370
32. Volberg L, Kulka M, Sust C, Lazarus H (2006) Speech intelligibility and the subjective assessment of speech quality in near real communication conditions. *Acta Acust United Acust* 92(3):406–416
33. Raake A, Egger S (2014) Quality and quality of experience. In: Möller S, Raake A (eds) *Quality of experience*. Springer International Publishing, pp 11–33
34. Kreiman J, Gerratt BR, Kempster GB, Erman A, Berke GS (1993) Perceptual evaluation of voice quality—review, tutorial, and a framework for future research. *J Speech Lang Hear Res* 36:21–40
35. ITU-T Recommendation P.800 (1996) *Methods for subjective determination of transmission quality*. International Telecommunication Union, Geneva
36. Qualinet (2013) *Qualinet white paper on definitions of quality of experience*. In: Le Callet P, Möller S, Perkins A eds (Version 1.2) *European network on quality of experience in multimedia systems and services (COST Action IC 1003)*, Lausanne, Switzerland. http://www.qualinet.eu/images/stories/QoE_whitepaper_v1.2.pdf
37. Geerts D, Moor KD, Ketyko I, Jacobs A, den Bergh JV, Joseph W, Martens L, Marez LD (2010) Linking an integrated framework with appropriate methods for measuring QoE. In: 2010 Second international workshop on quality of multimedia experience (QoMEX). Institute of Electrical and Electronics Engineers (IEEE)
38. ITU-T Recommendation E.800 (2008) *Definitions of terms related to quality of service*. International Telecommunication Union, Geneva
39. ITU-T Recommendation E.802 (2007) *Framework and methodologies for the determination and application of QoS parameters*. International Telecommunication Union, Geneva
40. Carroll JD (1972) Individual preferences and multidimensional scaling. In: Shepard RN, Romney AK, Nerlove S (eds) *Multidimensional scaling: theory and applications in the behavioral sciences*, vol. 1, pp 105–155
41. *Oxford Dictionary of English* (2010) Oxford University Press (OUP), Oxford
42. ITU-T Temporary Document TD 650rev1 (GEN/12) (2011) *Requirement specifications for P.TCA (Technical Cause Analysis)*. International Telecommunication Union, Geneva; Rapporteur Q.16/12 (L. Malfait), 2011
43. Möller S, Köster F, Skowronek J, Schiffner F (2013) Analyzing technical causes and perceptual dimensions for diagnosing the quality of transmitted speech. In: *Proceedings of the 4th international workshop on perceptual quality of systems (PQS 2013)*, pp 30–35

44. Köster F, Schiffner F, Guse D, Ahrens J, Skowronek J, Möller S (2015) Towards a MATLAB toolbox for imposing speech signal impairments following the P.TCA schema. *Audio Engineering Society Convention 139*, New York, pp 1–8
45. Möller S, Heute U (2012) Dimension-based diagnostic prediction of speech quality. In: *Proceedings of speech communication 10*, VDE, ITG Symposium, pp 1–4
46. Blauert J, Jekosch U (2003) Concepts of sound quality: some basic considerations. In: *INTER-NOISE 2003*, Jeju, Korea, pp 72–79
47. Blauert J (1996) *Spatial hearing. The Psychophysics of Human Sound Localization*. The MIT Press, Cambridge. Revised Edition
48. Osgood CE (1952) The nature and measurement of meaning. *Psychol Bull* 49(3):197–237
49. Letowski T (1989) Sound quality assessment: concepts and criteria. *J Audio Eng Soc* 37:1062
50. Quackenbush S, Barnwell T, Clements M (1988) Objective measures of speech quality. *Ser. Ellis Horwood Series in Artificial Intelligence*. Prentice Hall PTR, Upper Saddle River
51. Osgood C (1957) *The measurement of meaning*. University of Illinois Press, Urbana
52. Voiers W (1977) Diagnostic acceptability measure for speech communication systems. In: *Proceedings of the IEEE ICASSP*, Hartford, USA
53. Borg I, Groenen P (2005) *Modern multidimensional scaling—theory and applications*, 2nd edn. Springer Series in Statistics, New York
54. Möller S (2010) *Quality engineering*. Springer, Berlin
55. Poulton EC (1979) Models for biases in judging sensory magnitude. *Psychol Bull* 86(4):777
56. Zielinski S, Rumsey F, Bech S (2008) On some biases encountered in modern audio quality listening tests—a review. *J Audio Eng Soc* 56(6):427–451
57. Köster F, Guse D, Wältermann M, Möller S (2015) Comparison between the discrete ACR scale and an extended continuous scale for the quality assessment of transmitted speech. In: *Fortschritte der Akustik, DAGA 2015: Plenarvortr. u. Fachbeitr. d. 41. Dtsch. Jahrestg. f. Akust., DEGA*
58. *ITU-T Handbook on Telephonometry* (1992). International Telecommunication Union, Geneva
59. Guski R, Blauert J (2009) Psychoacoustics without psychology. In: *Proceedings of the NAG/DAGA 2009 international conference on acoustics*, vol. 3, pp 1550–1551
60. Reiter U, Brunnström K, Moor KD, Larabi M-C, Pereira M-C, Pinheiro A, You J, Zgank A (2014) Factors influencing quality of experience. In: Möller S, Raake A (eds) *Quality of experience*. Springer International Publishing, pp 55–72
61. Wolf S, Dvorak C, Kubichek R, South C, Schaphorst R, Voran S (1991) Future work relating objective and subjective telecommunications system performance. In: *Proceedings of the global telecommunications conference (GLOBECOM'91)*, pp 2129–2134
62. Möller S, Chan W-Y, Côté N, Falk TH, Raake A, Wältermann M (2011) Speech quality estimation: models and trends. *IEEE Sig Process Mag* 28(6):18–28
63. Takahashi A, Yoshino H, Kitawaki N (2004) Perceptual QoS assessment technologies for VoIP. *IEEE Commun Mag* 42(7):28–34
64. *ITU-T Recommendation P.564* (2007) Conformance testing for voice over IP transmission quality assessment models. International Telecommunication Union, Geneva
65. Hammer F (2006) *Quality aspects of packet-based interactive speech communication*, PhD thesis, Graz University of Technology
66. *ITU-T Recommendation P.59* (1993) Artificial conversational speech. International Telecommunication Union, Geneva
67. Guéguin M, Bouquin-Jeannès RL, Gautier-Turbin V, Faucon G, Barriac V (2008) On the evaluation of the conversational speech quality in telecommunications. *EURASIP J Adv Sig Proc* 2008
68. *ITU-T Recommendation P.48* (1988) Specification for an intermediate reference system. International Telecommunication Union, Geneva
69. *ITU-T Recommendation P.830* (1996) Subjective performance assessment of telephone-band and wideband digital codecs. International Telecommunication Union, Geneva

70. Appel R, Beerends J (2002) On the quality of hearing one's own voice. *J Audio Eng Soc* 50(4):237–248
71. Ohm J, Lüke H (2010) *Signalübertragung: Grundlagen der digitalen und analogen Nachrichtenübertragungssysteme*. Ser. Springer-Lehrbuch. Springer, Berlin
72. Deville Y, Fleurenceau L, Rech O, Audio G (2000) Telephone having automatic gain control means. US Patent 6,094,481
73. ITU-T Recommendation G.729 Annex B (1996) A silence compression scheme for G.729 optimized for terminals conforming to Recommendation V.70. International Telecommunication Union, Geneva
74. Ramirez J, Gorriz JM, Segura JC (2007) Voice activity detection. Fundamentals and speech recognition system robustness. InTech
75. ITU-T Recommendation O.41 (1994) Psophometer for use on telephone-type circuits. International Telecommunication Union, Geneva
76. ITU-T Recommendation P.810 (1996) Modulated noise reference unit (MNRU). International Telecommunication Union, Geneva
77. Benesty J, Chen J, Huang Y, Cohen I (2009) *Noise reduction in speech processing*. Ser. Springer Topics in Signal Processing. Springer, Berlin
78. Davis G (2002) *Noise reduction in speech applications*. Ser. Electrical Engineering & Applied Signal Processing Series. CRC Press, Boca Raton
79. ITU-T Recommendation P.835 (2003) Subjective test methodology for evaluating speech communication systems that include noise suppression algorithm. International Telecommunication Union, Geneva
80. Vary P, Martin R (2006) *Digital speech transmission: enhancement, coding and error concealment*. Wiley, Chichester
81. ITU-T Recommendation G.711 (1988) Pulse code modulation (PCM) of voice frequencies. International Telecommunication Union, Geneva
82. ITU-T Recommendation G.726 (1990) 40, 32, 24, 16 kbit/s adaptive differential pulse code modulation (ADPCM). International Telecommunication Union, Geneva
83. Tremain ET (1982) The government standard linear predictive coding algorithm: LPC10
84. Welch VC, Tremain TE, Campbell JP (1989) A comparison of US Government standard voice coders. In: Military communications conference, 1989, MILCOM '89, conference record on bridging the gap. interoperability, survivability, security, vol 1. IEEE, pp 269–273
85. ITU-T Recommendation P.729.1 (2006) G.729-based embedded variable bit-rate coder: an 8–32 kbit/s scalable wideband coder bitstream interoperable with G.729. International Telecommunication Union, Geneva
86. Ahson S, Ilyas M (2008) *VoIP handbook: applications, technologies, reliability, and security*. CRC Press, Boca Raton
87. ITU-T Recommendation G.191 (2010) Software tools for speech and audio coding standardization. International Telecommunication Union, Geneva
88. Perkins C (2003) *RTP: audio and video for the internet*, 1st edn. Addison-Wesley Professional, Boston
89. Ogunfunmi T, Togneri R, Narasimha M (eds) *Speech and audio processing for coding, enhancement and recognition*. Springer Science and Business Media
90. Gabrielsson A, Sjögren H (1979) Perceived sound quality of sound-reproducing systems. *J Acoust Soc Am* 65(4):1019–1033
91. Bernex E, Barriac V (2002) Architecture of non-intrusive perceived voice quality assessment. In: Proceedings of the international conference on measurement of speech and audio quality in networks (MESAQIN), CZ-Prague
92. Etame T, Faucon G, Gros L, Le Bouquin Jeannes R, Quinquis C (2008) Characterization of the multidimensional perceptive space for current speech and sound codecs. In: Proceedings of the 124rd AES convention, no 7410, NE–Amsterdam
93. McDermott BJ (1969) Multidimensional analyses of circuit quality judgements. *J Acoust Soc Am* 3(45):774–781

94. Hall JL (2001) Application of multidimensional scaling to subjective evaluation of coded speech. *J Acoust Soc Am* 110(4):2167
95. Mattila VV (2002) Ideal point modelling of speech quality in mobile communications based on multidimensional scaling (MDS). In: Proceedings of the 112th AES convention, no 5546. Audio Engineering Society, Munich
96. Mattila VV (2002) Descriptive analysis and ideal point modelling of speech quality in mobile communication. In: Proceedings of the 113th AES convention, no 5704. Audio Engineering Society, Los Angeles
97. Sen D (2001) Determining the dimensions of speech quality from PCA and MDS analysis of the diagnostic acceptability measure. MESAQUIN, CZ-Prague
98. Sen D (2004) Predicting foreground SH, SL and BNH DAM scores for multidimensional objective measure of speech quality. In: IEEE international conference on acoustics, speech, and signal processing. Institute of Electrical and Electronics Engineers (IEEE)
99. Wältermann M, Raake A, Möller S (2010) Quality dimensions of narrowband and wideband speech transmission. *Acta Acust United Acust*:1090–1103
100. Côté N, Gautier-Turbin V, Möller S (2007) Influence of loudness level on the overall quality of transmitted speech. In: Proceedings of the 123rd AES convention, no 7175, USA–New York, NY
101. ITU-T Recommendation P.10/G.100 (2006) Vocabulary for performance and quality of service. International Telecommunication Union, Geneva
102. ITU-T Recommendation P.806 (2014) A subjective quality test methodology using multiple rating scales. International Telecommunication Union, Geneva
103. ITU-T Contribution SG12-C228 (2011) P.MULTI subjective testing methodology number of relevant perceptual quality scales and additional validation of the test methodology. International Telecommunication Union, Geneva
104. ITU-T Recommendation P.862 (2001) Perceptual evaluation of speech quality (PESQ): an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs. International Telecommunication Union, Geneva
105. ITU-T Recommendation P.862.2 (2007) Wideband extension to recommendation P.862 for the assessment of wideband telephone networks and speech codecs. International Telecommunication Union, Geneva
106. ITU-T Recommendation P.863 (2011) Perceptual objective listening quality assessment. International Telecommunication Union, Geneva
107. Beerends JG, Stemerink JA (1994) A perceptual speech-quality measure based on a psychoacoustic sound representation. *J Audio Eng Soc* 42(3):115–123
108. Heute U, Möller S, Raake A, Scholz K, Wältermann M (2005) Integral and diagnostic speech-quality measurement: state of the art, problems, and new approaches. In: Proceedings of HO-Budapest, forum acousticum, pp 1695–1700
109. Scholz K (2008) Instrumentelle Qualitätsbeurteilung von Telefonbandsprache beruhend auf Qualitätsattributen. Shaker Verlag, Kiel
110. Huo L (2015) Attribute-based speech quality assessment—narrowband and wideband. Shaker Verlag, Kiel
111. Sen D, Lu W (2012) Objective evaluation of speech signal quality by the prediction of multiple foreground diagnostic acceptability measure attributes. *J Acoust Soc Am* 131(5):4087
112. ITU-T Temporary Document TD 438rev1 (GEN/12) (2014) Requirement specifications for P.AMD (perceptual approaches for multi-dimensional analysis). International Telecommunication Union, Geneva. Rapporteur Q.9/12 (J. Berger)
113. ITU-T Recommendation P.563 (2004) Single-ended method for objective speech quality assessment in narrow-band telephony applications. International Telecommunication Union, Geneva
114. Kim DS, Tarraf A (2004) Perceptual model for non-intrusive speech quality assessment. In: Proceedings of the IEEE ICASSP, Montreal, Canada
115. ITU-T Temporary Document TD (GEN/14) (2014) Technical requirement specification P.SPELQ. International Telecommunication Union, Geneva. Rapporteur Q.9/12 (J. Berger)

116. ITU-T Contribution SG12-C387 (2016) Proposal for a new work item on diagnostic single-ended quality indicators. International Telecommunication Union, Geneva
117. Köster F, Mittag G, Polzehl T, Möller S (2016) Non-intrusive estimation of noisiness as a perceptual quality dimension of transmitted speech. In: Proceedings of the 5th international workshop on perceptual quality of systems (PQS 2016), pp 74 – 78
118. Mittag G, Köster F, Möller S (2016) Non-intrusive estimation of the perceptual dimension coloration. In: Fortschritte der Akustik, DAGA 2016: Plenarvortr. u. Fachbeitr. d. 42. Dtsch. Jahrestg. f. Akust., DEGA
119. Köster F, Cercos-Llombart V, Mittag G, Möller S (2016) Non-intrusive estimation model for the speech-quality dimension loudness. In: Informationstechnische gesellschaft im VDE (ITG) conference on speech communication, 2016. ITG, pp 175–179
120. Cavanaugh JR, Hatch RW, Sullivan JL (1976) Models for the subjective effects of loss, noise, and talker echo on telephone connections. *Bell Syst Tech J* 55(9):1319–1371
121. Osaka N, Kakehi K (1986) Objective model for evaluating telephone transmission performance. *Rev Electr Commun Lab* 34(4):437–444
122. Johannesson N (1997) The ETSI computation model: a tool for transmission planning of telephone networks. *IEEE Commun Mag* 35(1):70–79
123. Zollinger SA, Brumm H (2011) The lombard effect. *Curr Biol* 21(16):R614–R615
124. Lombard E (1911) Le signe de l'élévation de la voix. *Annales des Maladies de L'Oreille et du Larynx* 37:101–119
125. Lane H, Tranel B (1971) The lombard sign and the role of hearing in speech. *J Speech Lang Hear Res* 14(4):677
126. Proakis J (2007) *Digital signal processing: principles, algorithms, and applications*, 4/E. Pearson Education, New York
127. Vljaj D, Kacic Z (2011) The influence of lombard effect on speech recognition. In: *Speech Technologies. InTech*
128. ITU-T Recommendation P.831 (1998) Subjective performance evaluation of network echo cancellers. International Telecommunication Union, Geneva
129. ITU-T Recommendation G.107 (2011) The E-model: a computational model for use in transmission planning. International Telecommunication Union, Geneva
130. Egger S, Reichl P, Schoenenberg K (2014) Quality of experience and interactivity. In: Möller S, Raake A (eds) *Quality of experience*. Springer International Publishing, pp 149–161
131. Möller S, Engelbrecht K-P, Kühnel C, Naumann A, Wechsung I, Weiss B (2010) Evaluation of multimodal interfaces for ambient intelligence. In: *Human-centric interfaces for ambient intelligence*. Elsevier BV, pp 347–370
132. Möller S, Engelbrecht K-P, Kühnel C, Wechsung I, Weiss B (2009) A taxonomy of quality of service and quality of experience of multimodal human-machine interaction. In: *International workshop on quality of multimedia experience, QoMEX 2009*. IEEE, pp 7–12
133. ITU-T Recommendation G.114 (2003) One-way transmission time. International Telecommunication Union, Geneva
134. ITU-T Recommendation P.805 (2007) Subjective evaluation of conversational quality. International Telecommunication Union, Geneva
135. ITU-T Contribution SG12-C35 (1997) Development of scenarios for short a conversation test. International Telecommunication Union, Geneva
136. Kitawaki N, Itoh K (1991) Pure delay effects on speech quality in telecommunications. *IEEE J Sel Areas Commun* 9(4):586–593
137. ITU-T Contribution SG12-C55 (2003) A subjective/objective test protocol for determining the conversational quality of a voice link. International Telecommunication Union, Geneva
138. ITU-T Temporal Document TD 27 (WP 2/12) (2005) Proposed scope for P.CQO. International Telecommunication Union, Geneva
139. ITU-T Temporary Document TD 965 Rev.1 (GEN/12) (2016) Revised status report of question 7/12. International Telecommunication Union, Geneva. Rapporteur Q.7/12 (P. Usai)
140. ETSI ETR 250 (1996) *Speech communication quality from mouth to ear for 3,1 kHz handset telephony across networks*. European Telecommunications Standards Institute, Sophia Antipolis

141. ITU-T Recommendation G.107.1 (2011) Wideband E-model. International Telecommunication Union, Geneva
142. ITU-T Recommendation G.113 (2007) Transmission impairments due to speech processing. International Telecommunication Union, Geneva
143. Hammer F, Reichl P, Raake A (2005) The well-tempered conversation: interactivity, delay and perceptual VoIP quality. In: Proceedings of the IEEE international conference on communications (ICC), Seoul, Korea
144. Köster F, Möller S (2014) Analyzing perceptual dimensions of conversational speech quality. In: Proceedings of 15th Annual Conference of the International Speech Communication Association (Interspeech 2014). ISCA Interspeech 2014 Proceedings, Singapore, pp 2041–2045
145. Kruskal J, Wish M (1978) Multidimensional scaling, quantitative applications in the social sciences. SAGE Publications, Newbury
146. Choe B (2001) Nonmetric multidimensional scaling of complex sounds: dimensions of preference ratings and perceived similarity of vehicle noises. Shaker, Kiel
147. Kruskal J (1964) Nonmetric multidimensional scaling: a numerical method. *Psychometrika* 29(2):115–129
148. Takane Y, Young FW, De Leeuw J (1977) Nonmetric individual differences multidimensional scaling: an alternating least squares method with optimal scaling features. *Psychometrika* 42:7–67
149. Kaiser HF (1958) The varimax criterion for analytic rotation in factor analysis. *Psychometrika* 23(3):187–200
150. MATLAB/Simulink (2014) Version 8.3.0.532 (R2014a). The MathWorks Inc., Natick
151. Bortz J (2005) Statistik. Springer, Berlin
152. Hinterleitner F, Norrenbrock C, Möller S, Heute U (2012) What makes this voice sound so bad? A multidimensional analysis of state-of-the-art text-to-speech systems. In: Proceedings of the 2012 IEEE workshop on spoken language technology (SLT), Miami, USA, pp 240–245
153. Puckette M (2015) Puredata. <http://puredata.info/>. Accessed 26 Aug 2015
154. The Telephone. <https://github.com/TheTelephone>. Accessed 19 Sep 2016
155. Helder GK (1966) Customer evaluation of telephone circuits with delay. *Bell Syst Tech J* 45(7):1157–1191
156. ITU-T Contribution SG12-C388 (2016) Proposal for a draft new recommendation for a subjective test methodology to analyse transmitted speech in a telephone conversation as a baseline for P.CQO. International Telecommunication Union, Geneva
157. Köster F, Möller S (2016) Introducing a new test-method for diagnostic speech quality assessment in a conversational situation. In: Fortschritte der Akustik—DAGA 2016: Plenarvortrag. u. Fachbeitr. d. 42. Dtsch. Jahrestg. f. Akust. DEGA, Berlin
158. ITU-T Contribution SG12-C304 (2016) Introducing a new subjective test methodology to analyse conversational quality. International Telecommunication Union, Geneva
159. Köster F, Möller S (2015) Perceptual speech quality dimensions in a conversational situation. In: Proceedings of the 16th Annual Conference of the International Speech Communication Association (Interspeech 2015). ISCA Interspeech 2015 Proceedings, Dresden, Germany, pp 2544–2548
160. Berger J, Llagostera A (2015) Multidimensional evaluation and predicting overall speech quality. In: Proceedings of 16th Annual Conference of the International Speech Communication Association (Interspeech 2015). ISCA Interspeech 2015 Proceedings, Dresden, Germany, pp 2549–2552
161. Belmudez B, Lewcio B, Möller S (2013) Call quality prediction for audiovisual time-varying impairments using simulated conversational structures. *Acta Acust United Acust* 99(5):792–805
162. Field A (2013) Discovering statistics using IBM SPSS statistics, 4th edn. SAGE Publications Ltd., Newbury

163. Miles J, Shevlin M (2001) *Applying regression and correlation: a guide for students and researchers*. SAGE Publications Ltd., 2001
164. F. Köster and S. Möller, "Analyzing the Relation Between Overall Quality and the Quality of Individual Phases in a Telephone Conversation," in *Proc. 17th Ann. Conf. of the Int. Speech Comm. Assoc. (Interspeech 2016)*. San Francisco, USA: ISCA Interspeech 2016 Proceedings, 2016, pp. 2493–2497
165. Glasberg BR, Moore BCJ (2002) A Model of Loudness Applicable to Time-Varying Sounds. *Journal of the Audio Engineering Society* 50(5):331–342