

Bibliography

- [1] Oulasvirta, Antti, Engelbrecht, Klaus, Möller, Sebastian & Jameson, Anthony. Internal report “*Error coding for freewoz data*”, MEMO Project, Deutsche Telekom Laboratories, 2006.
- [2] Bishop, Christopher M., *Neural Networks for Pattern Recognition*. Oxford University Press, 1995.
- [3] Möller, S., *Quality of telephone-based spoken dialogue systems*. Springer, New York, NY, 2005.
- [4] Walker, M.A., . Litman, D.J., Kamm, C.A., Abella, A., PARADISE: A framework for evaluating spoken dialogue agents. In: Proc. ACL/EACL 35th Meeting, Madrid, 271-280, 1997.
- [5] Walker, M.A., . Litman, D.J., Kamm, C.A., Abella, A., *Evaluating Spoken Dialogue Agents with PARADISE: Two Case Studies*. Computer Speech and Language 12, 1998.
- [6] Skowronek, J.. *Entwicklung von Modellierungsansätzen zur Vorhersage der Dienstqualität bei der Interaktion mit einem natürlichsprachlichen Dialogsystem*. Unpublished Diploma thesis, IKA, Ruhr-Universität, Bochum, 2002.
- [7] Möller, S. *Assessment and prediction of speech quality in Telecommunications*. Lower Academic Publ. Boston, USA, 2000.
- [8] Möller, S.. *A new taxonomy for the quality of telephone services based on spoken dialogue systems*. Proc. 3rd SIGDial Workshop on discourse and dialogue, 142-153. Philadelphia, USA, 2002.
- [9] ITU-T Rec. P.851: *Subjective quality evaluation of telephone services based on spoken dialogue systems*. International Telecommunication Union. Geneva, 2003.
- [10] ITU-T Suppl. 24 to P-Series Rec.: *Variables describing the interaction with spoken dialogue systems*. International Telecommunication Union. Geneva, 2003.
- [11] IST Project INSPIRE: Infotainment management with Speech Interaction via Remote-microphones and telephone interfaces. IST-2001-32746, <http://www.knowledge-speech.gr/inspire-project/index.html>.
- [12] A. Bühl, P. Zöfel. SPSS Version 11. *Einführung in die moderne Datenanalyse unter Windows*, 8. Auflage. Pearson Studium Addison-Wesley, 2002.
- [13] Weisberg, S. *Applied Linear Regression*, 2nd ed., John Wiley, New York, 1985.

- [14] Ostrom, C.W., Jr., *Time Series Analysis, Regression Techniques*, Second Edition, v.07-009: Newsbury Park, Sage Publications, 1990.
- [15] Larsen, L.B., *Combining objective and subjective data in evaluation of spoken dialogues*. Proc. ESCA Tutorial and Research workshop on Interactive Dialogue in multi-modal Systems. ESCA, 89-92. Bonn, D, 1999.
- [16] Stone, M. *Cross-validation: A review*. Statistics: A Journal of Theoretical and Applied Statistics, Volume 9, Issue 1, 127-139 (Mathematische Operationsforschung Statistik), 1978.
- [17] Wahba, G. and Wold, S. *A completely automatic French curve : Fitting spline functions by cross-validation*. Communication in Statistics – Theory and Methods, Volume 4, Issue 1, 1-17, 1975.

Appendixes

Tables

Correlations – BoRIS System

Variable	Statistics	Question B0	Question B23	Mean Questions B
Question B0	Pearson Correlation	1	0,745848778 (Sig. 1,01576E-35)	0,835064167 (Sig. 5,61259E-52)
Question B23	Pearson Correlation	0,745848778 (Sig. 1,01576E-35)	1	0,76373171 (Sig. 9,936E-39)
Mean Questions B	Pearson Correlation	0,835064167 (Sig. 5,61259E-52)	0,76373171 (Sig. 9,93692E-39)	1
Z-Wert: DD	Pearson Correlation	-0,337494908 (Sig. 1,40089E-06)	-0,31694958 (Sig. 6,00664E-06)	-0,348089051 (Sig. 5,38084E-07)
Z-Wert: STD	Pearson Correlation	0,123118379 (Sig. 0,08639458)	0,103569671 (Sig. 0,148573475)	0,202048734 (Sig. 0,00440963)
Z-Wert: UTD	Pearson Correlation	0,082796601 (Sig. 0,249841222)	-0,06494114 (Sig. 0,365828117)	0,119626491 (Sig. 0,094057747)
Z-Wert: SRD	Pearson Correlation	0,093579049 (Sig. 0,193184515)	-0,0340505 (Sig. 0,63564596)	0,141737646 (Sig. 0,046948518)
Z-Wert: URD	Pearson Correlation	-0,023108816 (Sig. 0,748464136)	-0,16024716 (Sig. 0,024856246)	-0,105749082 (Sig. 0,139150)
Z-Wert: # Turns	Pearson Correlation	-0,34064333 (Sig. 1,10045E-06)	-0,31320709 (Sig. 7,82802E-06)	-0,366387613 (Sig. 1,189E-07)
Z-Wert: # System Turns	Pearson Correlation	-0,34064333 (Sig. 1,10045E-06)	-0,31320709 (Sig. 7,82802E-06)	-0,366387613 (Sig. 1,18979E-07)
Z-Wert: # User Turns	Pearson Correlation	-0,34064333 (Sig. 1,10045E-06)	-0,31320709 (Sig. 7,82802E-06)	-0,366387613 (Sig. 1,189E-07)
Z-Wert: # System Words	Pearson Correlation	-0,327380333 (Sig. 2,98849E-06)	-0,28717998 (Sig. 4,48224E-05)	-0,335281476 (Sig. 1,464E-06)
Z-Wert: # User Words	Pearson Correlation	-0,243771574 (Sig. 0,000594408)	-0,24855386 (Sig. 0,000443435)	-0,195353393 (Sig. 0,0059409)
Z-Wert: WPST	Pearson Correlation	0,201494309 (Sig. 0,00473316)	0,22531007 (Sig. 0,001497894)	0,293692179 (Sig. 2,80581E-05)
Z-Wert: WPUT	Pearson Correlation	0,137327355 (Sig. 0,055566136)	0,043529712 (Sig. 0,54464413)	0,248646258 (Sig. 0,000426366)
Z-Wert: # Sys. Questions	Pearson Correlation	-0,314952339 (Sig. 7,31014E-06)	-0,29843537 (Sig. 2,15117E-05)	-0,353245408 (Sig. 3,55 E-07)
Z-Wert: # User Questions	Pearson Correlation	-0,029501183 (Sig. 0,682244578)	-0,01811336 (Sig. 0,801052174)	-0,021418017 (Sig. 0,7651404)
Z-Wert: # ASR Rejections	Pearson Correlation	-0,037950037 (Sig. 0,598383808)	-0,0353909 (Sig. 0,622395854)	-0,035186938 (Sig. 0,6235103)
Z-Wert: # Sys. Err. Mess.	Pearson Correlation	-0,095001453 (Sig. 0,186473406)	-0,14911542 (Sig. 0,036984126)	-0,113895282 (Sig. 0,1110226)
Z-Wert: # Help Requests	Pearson Correlation	0,0030324 (Sig. 0,966440494)	0,01132032 (Sig. 0,874869873)	-0,019988227 (Sig. 0,7804054)
Z-Wert: # Cancel Attempts	Pearson Correlation	-0,102351011 (Sig. 0,154502357)	-0,09611548 (Sig. 0,180208153)	-0,081819345 (Sig. 0,2530391)
Z-Wert: # Barge-Ins	Pearson Correlation	-0,139367599 (Sig. 0,051998772)	-0,11349509 (Sig. 0,113218759)	-0,084634302 (Sig. 0,2370223)
Z-Wert: SCT	Pearson Correlation	-0,267516815 (Sig. 0,000156329)	-0,2386866 (Sig. 0,000754231)	-0,27831637 (Sig. 7,494E-05)
Z-Wert: UCT	Pearson Correlation	-0,277625034 (Sig. 8,515E-05)	-0,25948292 (Sig. 0,000240065)	-0,302950541 (Sig. 1,510E-05)
Z-Wert: SCR	Pearson Correlation	-0,119071936 (Sig. 0,097325018)	-0,07988056 (Sig. 0,26572816)	-0,086783744 (Sig. 0,2252806)

Z-Wert: UCR	Pearson Correlation	-0,103523591 (Sig. 0,149808991)	-0,07573964 (Sig. 0,291384663)	-0,097131993 (Sig. 0,1745114)
Z-Wert: IR	Pearson Correlation	0,050525611 (Sig. 0,63819793)	0,023324742 (Sig. 0,829216716)	0,022221664 (Sig. 0,8362436)
Z-Wert: CA: # AP	Pearson Correlation	-0,301402827 (Sig. 1,85389E-05)	-0,27490935 (Sig. 9,63855E-05)	-0,323252708 (Sig. 3,600E-06)
Z-Wert: CA: # IA	Pearson Correlation	-0,329287489 (Sig. 2,59599E-06)	-0,31187415 (Sig. 8,59497E-06)	-0,354119194 (Sig. 3,306E-07)
Z-Wert: CA: # TF	Pearson Correlation	.	.	.
Z-Wert: CA: # IC	Pearson Correlation	.	.	.
Z-Wert: CA: AP	Pearson Correlation	0,269904466 (Sig. 0,00013572)	0,221945927 (Sig. 0,001769274)	0,310261826 (Sig. 9,121E-06)
Z-Wert: CA: IA	Pearson Correlation	-0,241631247 (Sig. 0,00066633)	-0,2490666 (Sig. 0,00043111)	-0,267149882 (Sig. 0,0001476)
Z-Wert: CA: TF	Pearson Correlation	.	.	.
Z-Wert: CA: IC	Pearson Correlation	.	.	.
Z-Wert: PA:CO	Pearson Correlation	-0,321819349 (Sig. 4,48137E-06)	-0,28238057 (Sig. 6,073E-05)	-0,345305056 (Sig. 6,714E-07)
Z-Wert: PA:PA	Pearson Correlation	-0,012621256 (Sig. 0,860985271)	-0,07263012 (Sig. 0,311703247)	0,015070522 (Sig. 0,8335188)
Z-Wert: PA:FA	Pearson Correlation	-0,291597835 (Sig. 3,534E-05)	-0,28925966 (Sig. 3,92276E-05)	-0,319761732 (Sig. 4,642E-06)
Z-Wert: AN:CO	Pearson Correlation	0,195008126 (Sig. 0,438078198)	0,369690737 (Sig. 0,13105563)	0,239969602 (Sig. 0,3374938)
Z-Wert: AN:PA	Pearson Correlation	-0,379027666 (Sig. 0,120859757)	-0,29865414 (Sig. 0,22865433)	-0,251754078 (Sig. 0,3135625)
Z-Wert: AN:FA	Pearson Correlation	-0,249177049 (Sig. 0,318707484)	-0,38914814 (Sig. 0,110455065)	-0,26759691 (Sig. 0,2830268)
Z-Wert: AN:IC	Pearson Correlation	.	.	.
Z-Wert: DARPA Score	Pearson Correlation	0,366539791 (Sig. 0,134628243)	0,478641205 (Sig. 0,044489143)	0,341708584 (Sig. 0,165177)
Z-Wert: DARPA Modified Error	Pearson Correlation	-0,440724874 (Sig. 0,06715728)	-0,49842496 (Sig. 0,03526515)	-0,37096844 (Sig. 0,1296260)
Z-Wert: IC	Pearson Correlation	0,116109153 (Sig. 0,10691413)	0,142638611 (Sig. 0,046677268)	0,099930678 (Sig. 0,1634470)
Z-Wert: UA	Pearson Correlation	-0,020809886 (Sig. 0,772767175)	-0,00378736 (Sig. 0,957983674)	-0,04314613 (Sig. 0,5471665)
Z-Wert: NES	Pearson Correlation	0,1154998 (Sig. 0,107860781)	0,013586133 (Sig. 0,850094415)	0,157661209 (Sig. 0,0269214)
Z-Wert: WES	Pearson Correlation	-0,056636182 (Sig. 0,431616799)	-0,07926652 (Sig. 0,269431929)	-0,025032408 (Sig. 0,7269651)
Z-Wert: WER	Pearson Correlation	-0,062106098 (Sig. 0,388400542)	-0,10280127 (Sig. 0,151625487)	-0,030173717 (Sig. 0,6741924)
Z-Wert: WA	Pearson Correlation	0,117476532 (Sig. 0,101925024)	0,054149928 (Sig. 0,450967398)	0,107356661 (Sig. 0,1332065)
Z-Wert: NES (iso)	Pearson Correlation	0,055211302 (Sig. 0,443312238)	-0,00815581 (Sig. 0,909671504)	0,116897925 (Sig. 0,1018578)
Z-Wert: WES (iso)	Pearson Correlation	-0,074249897 (Sig. 0,302258613)	-0,12462981 (Sig. 0,081780183)	-0,05028371 (Sig. 0,4828549)
Z-Wert: WER (iso)	Pearson Correlation	-0,083019234 (Sig. 0,248565745)	-0,12851057 (Sig. 0,072639263)	-0,055804126 (Sig. 0,4360557)
Z-Wert: WA (iso)	Pearson Correlation	0,147797368 (Sig. 0,039212827)	0,062258929 (Sig. 0,386001655)	0,148543123 (Sig. 0,0372324)
Z-Wert: Kappa per Configuration	Pearson Correlation	0,116081767 (Sig. 0,106085705)	0,144159458 (Sig. 0,043814073)	0,115806933 (Sig. 0,1051161)
Z-Wert: Kappa per Dialogue	Pearson Correlation	0,098236065 (Sig. 0,171851282)	0,062268902 (Sig. 0,385925419)	0,064135162 (Sig. 0,3705874)
Z-Wert: TS ordinal	Pearson Correlation	0,16029949 (Sig. 0,025184236)	0,285158855 (Sig. 5,09725E-05)	0,250056712 (Sig. 0,0003943)
Z-Wert: TS binär	Pearson Correlation	0,102735145 (Sig. 0,152952769)	0,241893596 (Sig. 0,000636143)	0,174331698 (Sig. 0,0142823)
Z-Wert: Question B1 Binary	Pearson Correlation	0,566776787 (Sig. 8,54989E-18)	0,590529399 (Sig. 1,2758E-19)	0,614201638 (Sig. 1,313E-21)

Correlations – INSPIRE System

Variable	Statistics	overall quality	overall satisfied	Use again
overall quality	Pearson Correlation	1	,620 (** Sig. ,000)	,500 (** Sig. ,000)
task success	Pearson Correlation	,600 (** Sig. ,000)	,325 (** Sig. ,008)	,252 (* Sig. ,033)
overall satisfied	Pearson Correlation	,620 (** Sig. ,000)	1	,709 (** Sig. ,000)
use again	Pearson Correlation	,500 (** Sig. ,000)	,709 (** Sig. ,000)	1
Z-Wert: dialogue duration (ms)	Pearson Correlation	-,307 (* Sig. ,013)	-,116 (Sig. ,370)	-,027 (Sig. ,827)
Z-Wert: system turn duration (ms)	Pearson Correlation	,157 (Sig. ,211)	,398 (** Sig. ,001)	,232 (Sig. ,057)
Z-Wert: user turn duration (ms)	Pearson Correlation	,009 (Sig. ,946)	,143 (Sig. ,267)	,050 (Sig. ,686)
Z-Wert: system response delay (ms)	Pearson Correlation	-,028 (Sig. ,827)	,096 (Sig. ,458)	,050 (Sig. ,687)
Z-Wert: user response delay (ms)	Pearson Correlation	-,092 (Sig. ,468)	-,065 (Sig. ,617)	,031 (Sig. ,801)
Z-Wert: # turns	Pearson Correlation	-,173 (Sig. ,165)	-,061 (Sig. ,635)	-,046 (Sig. ,709)
Z-Wert: # system turns	Pearson Correlation	-,325 (** Sig. ,008)	-,241 (Sig. ,059)	-,150 (Sig. ,221)
Z-Wert: # user turns	Pearson Correlation	-,325 (** Sig. ,008)	-,241 (Sig. ,059)	-,150 (Sig. ,221)
Z-Wert: words per system turn	Pearson Correlation	,141 (Sig. ,263)	,297 (* Sig. ,019)	,141 (Sig. ,251)
Z-Wert: # system words	Pearson Correlation	-,235 (Sig. ,060)	-,123 (Sig. ,340)	-,090 (Sig. ,465)
Z-Wert: words per user turn	Pearson Correlation	-,043 (Sig. ,732)	,076 (Sig. ,556)	,072 (Sig. ,559)
Z-Wert: # user words	Pearson Correlation	-,187 (Sig. ,135)	-,054 (Sig. ,675)	-,012 (Sig. ,926)
Z-Wert: # ASR rejections	Pearson Correlation	,114 (Sig. ,366)	,166 (Sig. ,197)	,180 (Sig. ,142)
Z-Wert: # system questions	Pearson Correlation	-,331 (** Sig. ,007)	-,238 (Sig. ,063)	-,161 (Sig. ,190)
Z-Wert: # system error mess.	Pearson Correlation	-,119 (Sig. ,345)	,177 (Sig. ,169)	,013 (Sig. ,918)
Z-Wert: # system help mess.	Pearson Correlation	-,441 (** Sig. ,000)	-,453 (** Sig. ,000)	-,320 (** Sig. ,008)
Z-Wert: # user questions	Pearson Correlation	-,019 (Sig. ,882)	-,008 (Sig. ,953)	,047 (Sig. ,702)
Z-Wert: SA:CO	Pearson Correlation	,052 (Sig. ,680)	,031 (Sig. ,810)	,075 (Sig. ,545)
Z-Wert: SA:IC	Pearson Correlation	-,013 (Sig. ,915)	-,032 (Sig. ,803)	,042 (Sig. ,732)
Z-Wert: SA:PA	Pearson Correlation	-,041 (Sig. ,744)	-,044 (Sig. ,734)	,176 (Sig. ,151)
Z-Wert: SA:FA	Pearson Correlation	-,090 (Sig. ,476)	,100 (Sig. ,441)	-,175 (Sig. ,152)
Z-Wert: DARPA score	Pearson Correlation	,038 (Sig. ,764)	,060 (Sig. ,642)	,040 (Sig. ,746)

Z-Wert: DARPA mod. error	Pearson Correlation	-,053 (Sig. ,674)	-,045 (Sig. ,729)	-,074 (Sig. ,546)
Z-Wert: # help requests	Pearson Correlation	,125 (Sig. ,322)	,151 (Sig. ,241)	,153 (Sig. ,213)
Z-Wert: # syst. correction turns	Pearson Correlation	-,216 (Sig. ,084)	-,197 (Sig. ,124)	-,131 (Sig. ,285)
Z-Wert: % system correction turns	Pearson Correlation	-,114 (Sig. ,364)	-,118 (Sig. ,361)	-,086 (Sig. ,486)
Z-Wert: # user correction turns	Pearson Correlation	-,444 (** Sig. ,000)	-,367 (** Sig. ,003)	-,266 (* Sig. ,028)
Z-Wert: % user correction turns	Pearson Correlation	-,408 (** Sig. ,001)	-,309 (* Sig. ,015)	-,247 (* Sig. ,043)
Z-Wert: # cancel attempts	Pearson Correlation	-,085 (Sig. ,500)	,164 (Sig. ,202)	,149 (Sig. ,227)
Z-Wert: # barge-ins	Pearson Correlation	-,068 (Sig. ,590)	-,178 (Sig. ,167)	-,171 (Sig. ,164)
Z-Wert: CA:AP	Pearson Correlation	-,191 (Sig. ,128)	-,239 (Sig. ,062)	-,124 (Sig. ,315)
Z-Wert: CA:IA	Pearson Correlation	-,400 (** Sig. ,001)	-,135 (Sig. ,295)	-,133 (Sig. ,280)
Z-Wert: % CA:AP	Pearson Correlation	,257 (* Sig. ,039)	-,092 (Sig. ,476)	,057 (Sig. ,646)
Z-Wert: % CA:IA	Pearson Correlation	-,257 (* Sig. ,039)	,092 (Sig. ,476)	-,057 (Sig. ,646)
Z-Wert: PA:CO	Pearson Correlation	-,147 (Sig. ,243)	-,122 (Sig. ,343)	-,020 (Sig. ,868)
Z-Wert: PA:PA	Pearson Correlation	-,285 (* Sig. ,021)	-,277 (* Sig. ,030)	-,317 (** Sig. ,008)
Z-Wert: PA:IC	Pearson Correlation	-,385 (** Sig. ,002)	-,273 (* Sig. ,032)	-,191 (Sig. ,118)
Z-Wert: implicit recovery	Pearson Correlation	-,004 (Sig. ,973)	,235 (Sig. ,066)	,271 (* Sig. ,026)
Z-Wert: # AVP	Pearson Correlation	-,281 (* Sig. ,023)	-,147 (Sig. ,253)	-,117 (Sig. ,341)
Z-Wert: # AVP correct	Pearson Correlation	-,118 (Sig. ,349)	-,072 (Sig. ,580)	-,042 (Sig. ,735)
Z-Wert: # AVP substituted	Pearson Correlation	-,270 (* Sig. ,030)	-,058 (Sig. ,654)	-,215 (Sig. ,079)
Z-Wert: # AVP inserted	Pearson Correlation	-,173 (Sig. ,169)	-,107 (Sig. ,409)	,056 (Sig. ,652)
Z-Wert: # AVP deleted	Pearson Correlation	-,337 (** Sig. ,006)	-,196 (Sig. ,126)	-,180 (Sig. ,142)
Z-Wert: # AVP not set	Pearson Correlation	-,323 (** Sig. ,009)	-,247 (Sig. ,053)	-,151 (Sig. ,218)
Z-Wert: information content	Pearson Correlation	,379 (** Sig. ,002)	,216 (Sig. ,092)	,220 (Sig. ,071)
Z-Wert: underst. error rate	Pearson Correlation	,458 (** Sig. ,000)	,320 (* Sig. ,011)	,339 (** Sig. ,005)
Z-Wert: # words	Pearson Correlation	-,187 (Sig. ,135)	-,054 (Sig. ,675)	-,012 (Sig. ,926)
Z-Wert: # words error per utt.	Pearson Correlation	-,332 (** Sig. ,007)	-,474 (** Sig. ,000)	-,383 (** Sig. ,001)
Z-Wert: word error rate	Pearson Correlation	-,299 (* Sig. ,016)	-,486 (** Sig. ,000)	-,380 (** Sig. ,001)
Z-Wert: word accuracy	Pearson Correlation	,299 (* Sig. ,016)	,486 (** Sig. ,000)	,380 (** Sig. ,001)
Z-Wert: # words (iso)	Pearson Correlation	-,240 (Sig. ,055)	-,241 (Sig. ,060)	-,103 (Sig. ,402)
Z-Wert: # words correct (iso)	Pearson Correlation	-,183 (Sig. ,144)	-,153 (Sig. ,235)	-,010 (Sig. ,937)

Z-Wert: # words subst. (iso)	Pearson Correlation	-,221 (Sig. ,076)	-,372 (** Sig. ,003)	-,507 (** Sig. ,000)
Z-Wert: # words deleted (iso)	Pearson Correlation	-,342 (** Sig. ,005)	-,498 (** Sig. ,000)	-,483 (** Sig. ,000)
Z-Wert: # errors per utt. (iso)	Pearson Correlation	-,334 (** Sig. ,007)	-,512 (** Sig. ,000)	-,549 (** Sig. ,000)
Z-Wert: word error per utt. (iso)	Pearson Correlation	-,434 (** Sig. ,000)	-,540 (** Sig. ,000)	-,543 (** Sig. ,000)
Z-Wert: word error rate (iso)	Pearson Correlation	-,356 (** Sig. ,004)	-,519 (** Sig. ,000)	-,534 (** Sig. ,000)
Z-Wert: word accuracy (iso)	Pearson Correlation	,356 (** Sig. ,004)	,519 (** Sig. ,000)	,534 (** Sig. ,000)
Z-Wert: weighted task success	Pearson Correlation	,169 (Sig. ,177)	,070 (Sig. ,586)	-,027 (Sig. ,824)

Questionnaires

BoRIS (English version):

TS-No.: _____

Date: _____

BoRIS

Dear participant!

Thank you for taking your time for this experiment!

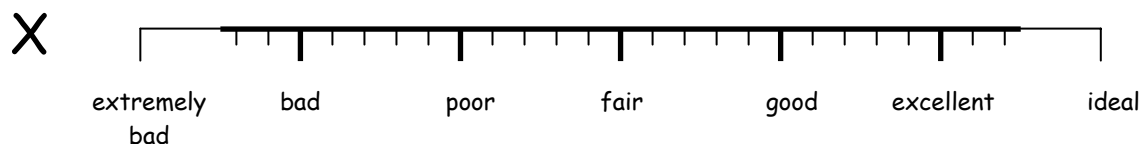
During the next hour you will get to know BoRIS via the telephone:
The Bochum restaurant information system.

This test will show how you experience a telephone call with BoRIS. For this aim, we ask you to call BoRIS five times. Before each call you will get a small task. At the end of **each** telephone call, we ask you to write down what you think about the system. You can do this easily by filling out a questionnaire.

Before the test starts, we would like to ask you to answer the questions given on the following pages. For the test evaluation, we need some personal information from your side, information which will be treated anonymously of course.

At the end of the **whole** experiment, we ask you to give an overall judgment about all calls you had with BoRIS.

For some assessments you will find the following scale:



Usually, your judgment should be in the range between bad and excellent. In case of an unpredictable extreme judgment, you can

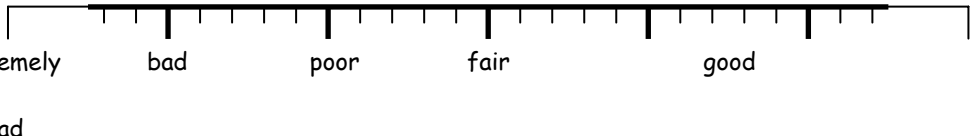
use the thinly drawn edges of the scale as well. Please also use the spaces between the grid marks, as depicted above.

Assess the system in a very self-confident way and remember during the whole test session:


Not **you** are tested, but **you** test our system!


And now: Have a lot of fun!


Part B


Overall impression: 

Information obtained from the system


1. The system provided the desired information: 


2. The provided information was... 

3. The information was... 

4. You would rate the information as... 

Communication with the system

5. How well did you feel understood by the system? 

6. You had to concentrate 

in order to understand
what the system expected
from you:

yes

no

7. How well was the system acoustically intelligible?
-

System behavior

8. You knew at each point of the dialogue what the system expected from you:
-

9. In your opinion, the system processed your specifications correctly:
-

10. The system's behavior was always as you expected:
-

11. How often did the system make mistakes?
-

12. The system reacted in the same way as humans do:
-

13. The system reacted...
-

14. You were able to control the dialogue in the desired way:
-


15. The system reacted...
-

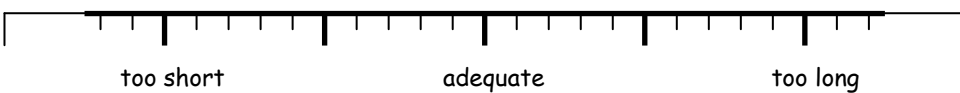
16. The system reacted in a ... way:
-


Dialogue

17. The system utterances were... 

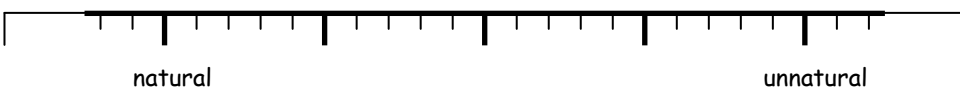
18. You perceived the dialogue as... 

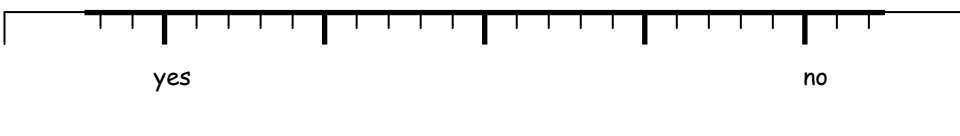
19. The course of the dialogue was... 

20. The dialogue was... 


21. The course of the dialogue was... 

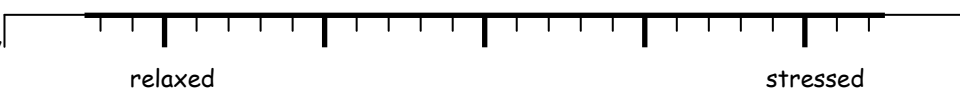
Your impression from the system

22. The system's voice was... 

23. Overall, you are satisfied with the dialogue: 

Personal impression

24. You perceived the dialogue as... 

25. During the dialogue, you felt ... 

INSPIRE System (german version):

VP-Nr.: _____

Datum: _____

INSPIRE-System

Liebe Teilnehmerin, lieber Teilnehmer!

Vielen Dank, dass Sie sich die Zeit für diesen Versuch nehmen.

Stellen Sie sich vor, dass Sie in einem innovativen Haushalt leben, in dem Sie verschiedene Hausgeräte durch Sprache steuern können: den Fernseher und den Videorekorder, den Anrufbeantworter, die Lampen, das Rollläden und den Ventilator. In diesem Versuch sollen Sie in drei Situationen mit diesen Geräten kommunizieren und bestimmte Aufgaben lösen. Diese Situationen werden Ihnen in Form kleiner Geschichten präsentiert, in denen die einzelnen Aufgaben **fett gedruckt** gekennzeichnet sind. Diese Aufgaben finden sich auch auf den Karteikarten auf dem Couchtisch.

Nach jeder Situation möchten wir von Ihnen erfahren, wie Sie die Qualität des Systems einschätzen. Dazu haben wir jeweils einen Fragebogen vorbereitet, den Sie bitte durch Ankreuzen ausfüllen. Beurteilen Sie selbstbewusst und bedenken Sie während des gesamten Versuchs:

Nicht **Sie** werden getestet, sondern **Sie** testen das System und bewerten es!

Uns ist Ihr spontanes Verhalten und Ihre offene Meinung über das System wichtig. Es ist auch nicht wichtig, wenn Sie einzelne Aufgaben trotz mehrerer Versuche nicht erfüllen können.

Die Dialoge die Sie mit dem System führen, werden aufgezeichnet, um das System verbessern zu können. Lassen Sie sich dadurch aber bitte nicht in Ihrem Verhalten oder in Ihrer Ausdrucksweise einschränken. Für die Auswertung des Tests benötigen wir einige Informationen über Sie, die selbstverständlich anonym bleiben. Deshalb beantworten Sie bitte zunächst die Fragen auf der folgenden Seite.

Nach den einleitenden Fragen haben wir beispielhaft einen Tag im Leben eines Ehepaars beschrieben, welches das INSPIRE-System benutzt. Lesen Sie sich diese Erzählung in Ruhe durch, um eine Vorstellung von den Möglichkeiten des Systems zu bekommen.

Und nun: Viel Spaß beim Versuch!

Beurteilung der Interaktion

1. Gesamteindruck der Interaktion mit dem *INSPIRE-System*:



2. Erreichen der gewünschten Ziele:

2.1 Das System tat nicht immer das, was ich wollte.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

2.2 Die vom System gelieferten Informationen waren klar und deutlich.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

2.3 Die gelieferten Informationen waren unvollständig.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

2.4 Mit dem System lassen sich die Hausgeräte effizient bedienen.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

2.5 Das System ist unzuverlässig.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

3. Verständigung mit dem System:

3.1 Ich fühle mich gut vom System verstanden.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

3.2 Ich wusste zu jeder Zeit, was ich dem System sagen konnte.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

3.3 Ich musste mich konzentrieren, um das System akustisch zu verstehen.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

3.4 Die Stimme des Systems klang natürlich.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

4. Verhalten des Systems:

4.1 Das System reagierte zu langsam.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

4.2 Das System ist freundlich.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

4.3 Das System reagierte nicht immer wie erwartet.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

4.4 Ich wusste nicht immer, was das System von mir verlangte.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

4.5 Das System machte viele Fehler.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

4.6 Ich konnte auftretende Fehler leicht beheben.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

4.7 Das System reagierte wie ein Mensch.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

4.8 Das System verhielt sich kooperativ.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

5. Gespräch:

5.1 Ich konnte leicht den Gesprächsfaden verlieren.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

5.2 Das Gespräch verlief holprig.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

5.3 Ich konnte das Gespräch wie gewünscht lenken.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

5.4 Das Gespräch war zu lang.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

5.5 Das Gespräch führte schnell zum gewünschten Ziel.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

5.6 Die Gesprächsanteile waren gleich verteilt zwischen mir und dem System.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

6. Persönliche Wirkung:

6.1 Die Interaktion mit dem System war angenehm.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

6.2 Ich fühlte mich entspannt.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

6.3 Ich musste mich sehr auf die Interaktion mit dem System konzentrieren.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

6.4 Die Interaktion hat Spaß gemacht.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

6.5 Ich bin insgesamt mit dem System zufrieden.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

7. Benutzbarkeit des Systems:

7.1 Das System lässt sich nur schwer bedienen.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

7.2 Die Benutzung des Systems lässt sich leicht erlernen.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

7.3 Die Bedienung der Hausgeräte durch Sprache war komfortabel.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

7.4 Das System ist zu unflexibel.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

7.5 Das System ist nicht hilfreich zur Bedienung der Hausgeräte.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

7.6 Ich würde die Hausgeräte lieber auf andere Weise bedienen.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

7.7 Ich würde das System in Zukunft wieder benutzen.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab

7.8 Die Benutzung des Systems hat sich gelohnt.

stimme stark zu stimme zu unentschieden lehne ab lehne stark ab