

ASPEKTE VON SYMPATHIE: GESPRÄCHSSTRUKTUR UND GESPRÄCHSANTEIL

Benjamin Weiss¹ und Katrin Schoenenberg²

¹ *Quality and Usability*, ² *Assessment of IP-based Applications*, TU Berlin
benjamin.weiss@tu-berlin.de

Kurzfassung: Drei-Personen-Audiokonferenzen mit hoher Übertragungsqualität wurden maschinell hinsichtlich individuellem Sprechverhalten ausgewertet: Also Sprechzeiten, -dauern, -pausen, Turnwechsel, etc. Abschließend bewertete Sympathie der Gesprächsteilnehmer untereinander ist zwar dominant von Sympathiebewertungen abhängig, die direkt vor dem Gespräch erhoben wurden; in Regressionsmodellen entfallen jedoch auf die 4 bzw. 5 Parametrisierungen von Gesprächsstrukturen und -anteile etwa die Hälfte der erklärten Varianz.

1 Einleitung

In Kommunikationssituationen nutzen Gesprächspartner zur Verfügung stehende Informationen, um Eigenschaften des Gesprächspartners zu schätzen; der sogenannte Attributionsprozess [4]. Dieses Bild vom Gesprächspartner wird unter anderem auf der Valenzdimension bewertet, also wie positiv oder negativ, auch zu interpretieren als wie sympathisch oder unsympathisch, uns jemand ist. Hierbei spielt non-verbales Verhalten eine bedeutende Rolle [10]. Bekannte Ursachen von Sympathie umfassen hierbei positive Reize durch die situative Interpretation des Verhaltens des Gegenübers, räumliche Nähe, Ähnlichkeit, komplementäre Bedürfnisse, sowie Schönheit [5], wobei sich Ähnlichkeit und insbesondere Verhalten in Gesprächen manifestieren. Nach dem Modell der vier Aspekte einer Nachricht können solche Informationen in Selbstoffenbarung, Sachaspekt, Appell sowie Beziehungsaspekt eingeteilt werden [17]. Demnach werden bei Sprechhandlungen linguistische, aber auch para-linguistische und extra-linguistische Informationen übermittelt.

Bei rein akustischer Kommunikation – beispielsweise über Telefon – können visuelle Eindrücke auf die Sympathiebewertungen ausgeschlossen werden. Jedoch werden auch akustische Informationen verwendet, um Eigenschaften des Gesprächspartners zu schätzen und daraufhin zu bewerten. Hierbei sind bereits einige para- und extralinguistische Faktoren untersucht worden:

Alter: Ältere Stimmen werden weniger angenehm bewertet als jüngere. Hierbei sind jüngere Erwachsene kritischer als ältere [2].

Geschlecht: Weibliche Sprecher werden weniger angenehm bzw. sympathisch bewertet als männliche [2, 21].

Attraktivität: Tiefe männliche Stimmen mit niedrigen Formantfrequenzen werden von Frauen angenehmer und attraktiver empfunden [3]. Allgemein werden attraktive Stimmen (wie auch attraktive Gesichter) positiv bewertet [23].

Persönlichkeit: Sprechern werden aufgrund von Stimme und Sprechweise Persönlichkeitsmerkmale zugeschrieben [19, 11]. Eine ähnliche Persönlichkeit wird hierbei vom Gesprächspartner bevorzugt [1].

Sprechstil: Allgemein werden tiefere Stimmlage, hohes Tempo, wenige Häsitationen und Versprecher sowie passende Intonation und Pausensetzung positiv bewertet [18, 13, 16]. Während eine hohe Sprecherkompetenz mit Sympathie in Beziehung steht [22] sind andere Faktoren noch nicht auf ihren Zusammenhang mit Sympathiebewertungen überprüft. Beispielsweise zeigt sich eine positive Korrelation zwischen Sprechtempo und Kompetenzzuschreibung, jedoch entspricht der Zusammenhang zwischen Sprechtempo und zugeschriebenem Wohlwollen einem Idealpunktmodell [8].

Solche akustischen Informationen betreffen wie Aussehen und Kleidung vornehmlich die Ebene der Selbstoffenbarung. Noch ist nicht klar, wie nachhaltig solche „vocal stereotypes“ [8] in tatsächlichen Gesprächen sind, da sie weitestgehend unter rein passiven Experimentalbedingungen untersucht werden. Daher stellt sich auch die Frage, welchen Einfluss die Ebenen des Appells oder der Beziehung auf die Sympathiebewertungen haben. Diese manifestieren sich beispielsweise anhand sprachlicher Handlungen im Gesprächs. Es gibt jedoch kaum Studien zum Einfluss individueller Gesprächsverläufe bei unbekanntem Gesprächspartnern auf Sympathiebewertungen. Statt dessen sind viele Untersuchungen aus der Sozialpsychologie auf mittel- und langfristige Einflussgrößen konzentriert, wie etwa räumliche Nähe, erfahrene Freundlichkeit oder gegenseitige (soziale, charakterliche, körperliche) Ähnlichkeit [1].

Im der vorliegenden Analyse wurde der Beziehungsaspekt anhand von Gesprächsverläufen untersucht. Hierzu wurden Drei-Personen-Telefonkonferenzen simuliert. Dieser Ansatz, Gesprächsstrukturen zu parametrisieren wurde bereits erfolgreich für die Vorhersage von Qualität von Sprachdialogsystemen verwendet [20, 12]. Nun soll der Einfluss solcher Parameter auf Sympathiebewertungen überprüft werden, wobei der erste Eindruck anhand von akustischen und visuellen Informationen kontrolliert wurde.

Zugeschriebene Sympathie wird hierbei als Resultat von vier Faktoren angesehen, die alle innerhalb der Kommunikationssituation interpretiert werden:

1. dem ersten (akustischen) Eindruck, wie angenehm, attraktiv und kompetent ein Sprecher klingt (Selbstoffenbarungsaspekt)
2. ob er/sie zu mir passt, bzw. mir ähnelt (Beziehungsaspekt)
3. welche soziale Beziehung er/sie signalisiert (Selbstoffenbarungsaspekt bzw. auf Hörerseite der Beziehungsaspekt)
4. was er/sie von mir möchte (Apellaspekt).

Gesprächsstrukturen bilden sich demnach besonders unter Punkt 3) und 4) ab, ggf. auch unter Punkt 2): Werden also ähnliche Gesprächsanteile, unterbrechende Teilnehmer oder Vielsprecher sympathischer Bewertet als andere? Ausgeschlossen wird hierbei der inhaltliche Aspekt; es wird dabei postuliert, dass (vorgegebene) Fakten bei sachlichen Gesprächen nicht die Bewertung der Gesprächsteilnehmer beeinflussen (Subtraktionsregel [5]).

2 Material

Die Gesprächsaufnahmen und Sympathiebewertungen, die im Rahmen dieser Arbeit analysiert werden, sind Teil einer Studie zur Qualitätswahrnehmung von mehr-Personen Telefonkonferenzen an den T-Labs, TU-Berlin. Insgesamt nahmen 36 Personen (9 Frauen / 30 Männer) in gleichgeschlechtlichen und gemischten Gruppen à drei an der Studie teil. Das mittlere Alter der Personen lag bei 36,2 Jahren (SD = 12,2), wobei alle Teilnehmer einander zuvor unbekannt waren. Eine Erfahrung in mindestens drei mehr-Personen Telefonkonferenzen im letzten Jahr wurde als Voraussetzung festgelegt, damit eine gewissen Bekanntheit mit der Technologie als gegeben betrachtet werden konnte. Im Mittel gaben die Teilnehmer 34.65 (SD = 75.50) Telefonkonferenzen im beruflichen und privaten Kontext an.

Nach Ankunft in den Versuchsräumen klärte der Versuchsleiter die teilnehmenden Personen über ihre Aufgabe für den nachfolgenden Versuch auf. Jede Person nahm im Anschluss Platz in einem der drei schallgeschützten Versuchsräume (nach ITU-T Rec. P.800 [7]). Die drei Räume waren durch ein Konferenzsystem, implementiert in Pd [14], verbunden. Das System stellte initial eine breitbandige Verbindung bereit, die eine Laustärkeabdämpfung von 23.1 dB SPL (Testsignal von 61.3 dB SPL, gemessen am fernen Ende mit 84.4 dB SPL) aufwies. Zur Kommunikation trugen die Versuchspersonen geschlossene Beyerdynamics DT 290 Headsets. Anhand der Mikrofonsignale der Headsets wurden alle Gespräche aufgenommen, um in einer späteren Analyse Maße der Gesprächsstruktur extrahieren zu können. Der Versuchsleiter initiierte das Gespräch der Teilnehmer durch eine Melodie.

Nach einem Eingewöhnungsgespräch in der bestmöglichen Qualität (ohne eingeführte Störungen) durchliefen die Teilnehmer weitere neun Konversationen, wobei ihnen die jeweilige technische Bedingung unbekannt blieb. Unter den neun technischen Bedingungen befand sich erneut eine Bedingung mit bestmöglicher Qualität. Die Szenarien zur Anleitung der Gespräche entstammten aus der Sammlung der 3 Konversations-Tests [15]. Hierbei handelt es sich um semi-strukturierte Aufgaben, die zu typischen natürlichen Geschäftsgesprächen mit Themen wie der Planung einer Konferenzstätte oder der Lieder einer neuen Musik CD führen. Jede Person erhält dabei pro Szenario einen unterschiedlichen Bogen mit Informationen zum Austausch, zur Klärung und zum Erfragen. Im diesem Versuch blieb das Thema des Eingewöhnungsszenarios konstant für alle Gruppen, die Thematik des zweiten Szenarios in bestmöglicher Qualität konnte jedoch variieren. Nach jeder Konversationen wurden die Teilnehmer gebeten die wahrgenommene Gesamtqualität der Verbindung, die empfundenen Aufmerksamkeit und die wahrgenommenen Sympathie gegenüber den anderen beiden Gesprächspartner auf Skalen zu bewerten.

Im Weiteren sollen nur die beiden Gespräche in bestmöglicher Qualität und die Sympathiebewertungen betrachtet werden. Die Sympathie wurde jeweils auf einer kontinuierlich Skala mit den Endpunkten „sehr sympathisch“ und „sehr unsympathisch“ getrennt für jeden der beiden anderen Gesprächspartner bewertet. Ein Sympathiewert von null steht für „sehr sympathisch“, ein Wert von 10 für „sehr unsympathisch“.

Segmentationen der einzelnen Audiokanäle wurden automatisch anhand eines Voice Activity Detection Algorithmus [9] durchgeführt, um Grenzen und Dauern der Sprechanteile für jeden Teilnehmer zu bestimmen. Mit Matlab Skripten wurden daraufhin, basierend auf einem Zustandsmodell [6], abgeleitete Werte extrahiert: Mittlere Dauern für Sprachabschnitte (Turns) und Pausen, Anteile von Sprachabschnitten an der Gesamtdauer und Pausen an der eigenen Sprechdauer, Dauern für Sprachabschnitte mit zwei und drei Sprechern, Art der Übergänge zwischen Sprechern (also mit einer Pause dazwischen oder als „Unterbrechung“).

3 Ergebnisse

Die hier dargestellten Ergebnisse beziehen sich auf die Bedingung mit optimaler Übertragungsqualität, also einer vollbandigen Telefonverbindung in ruhiger Umgebung ohne weitere technischen Codierungen oder Übertragungsfehler. Da die Teilnehmer gemeinsam begrüßt wurden, und sich somit kurz sehen konnten, werden die Sympathiewerte mit der Kontrollvariablen nach der kurzen Eingewöhnungsphase analysiert, um den Einfluss des Gesprächs von dem Eindruck vor dem Gespräch auf die finalen Beurteilungen trennen zu können. Für zwei Teilnehmer wurden keine Bewertungen abgegeben; von den finalen 37 Gesprächsteilnehmern sind für 10 jeweils nur eine anstatt zwei Bewertungen vorhanden. Im Fall von nur einer Bewertung wurde nur diese verwendet.

Zudem weist die Verteilung der für beide Bewerter gemittelten Werte keine Normalverteilung

auf, weshalb die beiden Daten > 4 ausgeschlossen wurden, um unrealistisch hohe Korrelationen zu vermeiden (siehe Abbildung 1).

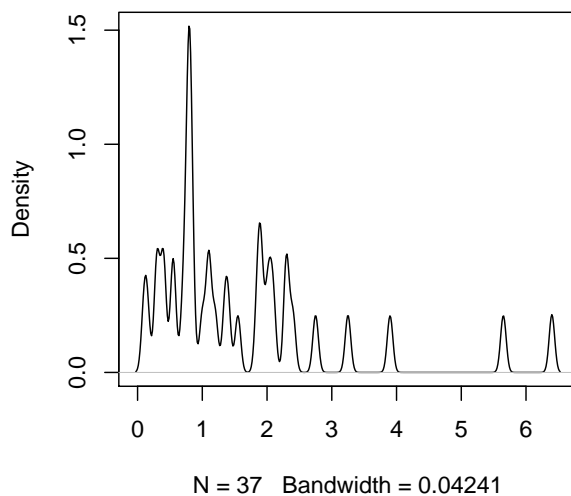


Abbildung 1 - Dichtefunktion der absoluten gemittelten Bewertungen. 10 Einzelwerte vorhanden.

Gemittelte Sympathiebewertungen zeigen einen signifikanten Zusammenhang mit automatisch extrahierten Parametern der Gespräche (multiple lineare Regression mit schrittweiser Parametereindebnung anhand des AI-Kriteriums): $R_{adj}^2 = .72$, RMSE= .49. Ohne die Kontrollvariable (Basisbewertung: Sympathiebewertung vor dem Gespräch) ist die Varianzaufklärung signifikant geringer ($R_{adj}^2 = .36$, RMSE= .73). Die Parameter sind in Tabelle 1 eingetragen, der Vergleich der geschätzten gegenüber den echten Bewertungen in Abbildung 2.

Tabelle 1 - Ergebnisse des linearen Modells für gemittelte Sympathiebewertungen.

Parameter	Beta	p-Wert
Basisbewertung	0.61	1.33e-06***
mittlere Verweilzeit in alleinigem Sprechen	-2.47	1.79e-06***
Anzahl eigener Sprechpausen im Turn	-0.03	6.72e-06***
Anzahl „Unterbrechungen“ anderer durch Sprecher / Sprechzeit andere	-0.19	0.0120*
Anzahl eigener Turnenden durch Pause / eigene Sprechzeit	-0.22	2.88e-07***
Anzahl aller Turnwechsel durch Pausen	0.01	0.0732.

Es konnte ein signifikanter Zusammenhang zwischen Parametrisierungen der Gespräche und Veränderungen von Sympathiebewertungen nach den Gesprächen festgestellt werden. Fünf Parameter wurden mit der Basisbewertung in einem einfachen linearen Modell erfasst. Diese alleinige Sprechzeiten, Verhalten im Turn und Turnwechsel:

Die Bewertungen sind negativer für längere Verweilzeiten in alleinigem Sprechen, mehr Pausen während des eigenen Turns, sowie für mehr Phasen, wo der Bewertete während des Turns eines anderen mit diesem gleichzeitig spricht, relativ zur Sprechzeit der anderen Teilnehmer. Der zuletzt genannte Parameter kann hierbei sowohl Rückmeldungen vom Hörer (back-channel zur Bestätigung und Motivation des Fortfahrens), als auch missglückte Versuche, das Rederecht zu übernehmen, umfassen. Bei den letzten beiden Parametern handelt es sich um Anzahl beendeter Turns mit Pause im Verhältnis zur eigenen Sprechzeit und die Gesamtanzahl aller Turnwechsel durch eine Pause anstatt von „Unterbrechungen“.

Während sich hier also die gemittelten Sympathiebewertungen anhand der Basisbewertung und des individuellen Gesprächsanteils mit 5 Parametern beschreiben lässt (Pearson's $r = .86$, bzw. $r = .66$ ohne Basisbewertung), ist es doch bemerkenswert, dass für die 27 Teilnehmer,

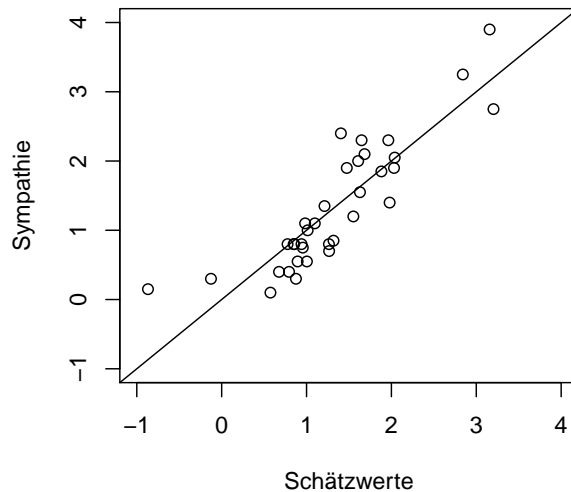


Abbildung 2 - Vergleich der Regressionswerte mit echten Bewertungen für 35 Teilnehmer.

für die beide Beurteilungen vorliegen, die Bewertungen recht stark streuen (vgl. Abbildung 3): Sortiert nach besseren und schlechteren Bewertungen jedes Paares ergibt sich eine Korrelation von $r = .51$ nach Ausschluss des einen Extremwertes.

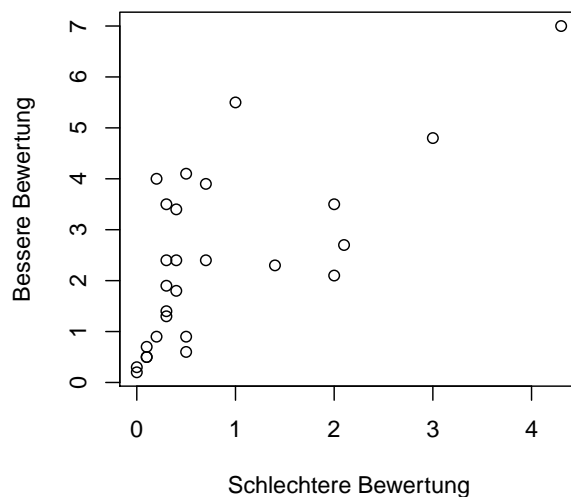


Abbildung 3 - Vergleich der beiden Bewertungen für 27 Teilnehmer.

Wird nun versucht, diese Einzelbewertungen zu modellieren, indem extrahierte Parameter des bewerteten Teilnehmers zu denen des Bewerter ins Verhältnis gesetzt werden, ergibt sich eine deutlich geringere Varianzaufklärung gegenüber den gemittelten Daten. ($R_{adj}^2 = .47$, RMSE= 1.05, ohne Basisbewertung $R_{adj}^2 = .25$, RMSE= 1.26), wie in Abbildung 4 zu sehen ist (vgl. auch Tabelle 2).

4 Diskussion

Nicht nur die ersten Beurteilungen der Gesprächspartner, sondern auch Parametrisierungen der Gesprächsverläufe korrelieren signifikant mit den finalen Sympathiewerten. Die 5 relevanten Parameter umfassen hierbei zahlreiche Parametergruppen: mittlere Dauern, Pausenverhalten, sowie Turnwechsel. Aufgrund der automatischen Extraktion sind diese Parameter ersteinmal deskriptiv zu verstehen. Da Teilnehmer negativ beurteilt werden, falls sie länger alleine sprechen und dabei mehr Pausen machen, kann vermutet werden, dass sich hier ein unkoopera-

Tabelle 2 - Ergebnisse des linearen Modells für ungemittelte Sympathiebewertungen. Daten entsprechen dem Verhältnis eines Parameters des bewerteten zu bewertenden Teilnehmers.

Parameter	Beta	p-Wert
Basisbewertung	0.54	3.03e-05***
Anzahl eigener Turnenden durch „Unterbrechungen“/ Anzahl eigener Turnenden durch Pause	0.61	0.0309*
eigener Sprechanteil an Gesamtdauer	-1.34	0.0155*
Anzahl „Unterbrechungen“ anderer durch Sprecher	-0.78	0.0864.
Anteil eigener Stille an Gesamtzeit	-2.22	0.1223

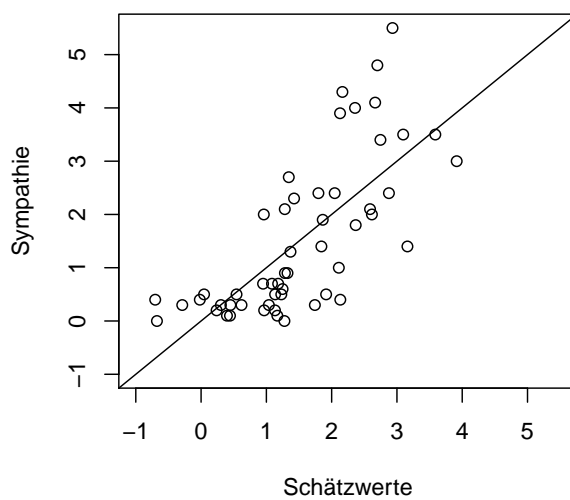


Abbildung 4 - Vergleich der Regressionswerte mit 53 echten Bewertungen für 27 Teilnehmer.

tives, uneffizientes, ggf. sogar irrelevantes Sprachverhalten negativ auswirkt. Auch der negative Zusammenhang mit der relativen Anzahl von Turnenden mit Pausen könnte darauf hinweisen, dass hier nicht aktiv zugehört wurde, sondern lediglich darauf gewartet, bis der Sprecher fertig ist. Bei der relativen Anzahl von „Unterbrechungen“, die nicht zu einem Turnwechsel führten wäre spannend, ob es sich bei diese tatsächlich vornehmlich um Unterbrechungen handelt, die mit negativen Beurteilungen korrelieren, da häufiges Backchannelling eher mit einer positiven Einstellung der Gesprächspartners in Bezug gesetzt wird. Jedenfalls tritt ein häufiges „Unterbrechen“, also gleichzeitig sprechen, mit negativeren Bewertungen des Unterbrechers auf. Der letzte Parameter lässt sich aufgrund seines absoluten Charakters nicht interpretieren, da er beispielsweise auch die reine Gesprächsdauer oder Anzahl Turnwechsel repräsentieren könnte.

Die jeweils beiden Einzelbewertungen sind jedoch wenig einheitlich. Es ist nicht gelungen, diese Einzelbewertungen anhand Parameterverhältnisse für Bewertenden und Bewertetem ähnlich gut zu beschreiben wie die gemittelten; wobei hier aufgrund der Fülle von Parametern auch weniger abgeleitete, also über Gesamtanzahl oder -dauer relativierte, verwendet wurden. Dennoch ist das Ergebnis für die Einzelwerte vielversprechend, mit der Aufbereitung des Korpus den Mehrwert der Drei-Personen-Gespräche – individuelles Erleben einen Teilnehmers im gleichen Gespräch – erarbeiten zu können.

5 Fazit

Interpersonelle Bewertung zuvor unbekannter Gesprächspartner steht nicht nur in Beziehung zu Stimme und Sprechweise, sondern auch zum Gesprächsverlauf. Damit erweitert diese Arbeit

bisherige Untersuchungen zur Sympathie von Stimme und Sprechweise unbekannter Personen auf tatsächliches Interaktionsverhalten. Mit der Festlegung auf Veränderungen der Sympathiebewertungen nach einem kurzen Training zu denen nach dem untersuchten Gespräch durch die Kontrollvariable, wurde zwar versucht, ursächliche Parameter zu identifizieren; jedoch kann hier bislang lediglich von einem signifikanten Zusammenhang zwischen Gesprächsparametern und interpersoneller Beurteilung gesprochen werden. Für den Nachweis eines ursächlichen Zusammenhanges bedarf es kontrolliert und systematisch veränderter Parameter, beispielsweise mit Regieanweisungen geschulter Sprecher oder spezieller Sprachdialogsysteme. Insbesondere kann der erste Eindruck vor dem analysierten Gesprächsteil wiederum das Gesprächsverhalten beeinflussen.

Derzeit wird das Korpus aufwändig nach Dialogakten händisch annotiert, um zum Einen die hier verwendete automatische Segmentation zu validieren, zum Anderen jedoch auch, um die relevanten Parameter nicht bloß zu präsentieren, sondern auch interpretieren zu können; auch um später den Appellaspekt zu erfassen.

6 Danksagung

Die hier vorgestellten Arbeiten wurden von der DFG gefördert (WE 5050/1-1).

Literatur

- [1] ARONSON, E., T. WILSON und R. M. AKERT: *Social Psychology*. Prentice Hall, 7 Aufl., 2009.
- [2] DEAL, L. und H. OYER: *Ratings of vocal pleasantness and the aging process*. *Folia Phoniatica*, 43:44–48, 1991.
- [3] FEINBERG, D., B. JONES, A. LITTLE, D. BURT und D. PERRETT: *Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices*. *Animal Behaviour*, 69:561–568, 2005.
- [4] HELFRICH, H.: *Paralinguistic Behaviors and Culture*. In: SPIELBERGER, C. D. (Hrsg.): *Encyclopedia of Applied Psychology*, Bd. 2, S. 797–813. Academic Press, Amsterdam, 2004.
- [5] HERKNER, W.: *Lehrbuch Sozialpsychologie*. Huber, Bern, 1991.
- [6] HOELDTKE, K. und A. RAAKE: *Conversation analysis of multi-party conferencing and its relation to perceived quality*. In: *Proc of the. Int. Conf. on Communications (ICC), IEEE*, S. 1–5. Kyoto, Japan, 2011.
- [7] ITU-T REC. P.800: *Methods for subjective determination of transmission quality*. International Telecommunication Union, Geneva, 1996.
- [8] KREIMAN, J. und D. VAN LANCKER SIDTIS: *Foundations of Voice Studies: An Interdisciplinary Approach to Voice Production and Perception*. Wiley-Blackwell, Chichester, 2011.
- [9] LUENGO, I., E. NAVAS, I. ODRIOZOLA, I. SARATXAGA, I. HERNAEZ, I. SAINZ und D. ERRO: *Modified LTSE-VAD algorithm for applications requiring reduced silence frame misclassification*. In: *Proc. of the LREC*, S. 1539–1544, 2010.

- [10] MEHRABIAN, A.: *Some referents and measures of nonverbal behavior*. Behavioral Research Methods and Instrumentation, 1:213–217, 1969.
- [11] MÖLLER, R.: *Stimme und Persönlichkeit – Attribution von Persönlichkeitseigenschaften eines Sprechers anhand von physikalischen Parametern der Stimme*. Franzbecker, Hildesheim, 2009.
- [12] MÖLLER, S., K.-P. ENGELBRECHT und R. SCHLEICHER: *Predicting the Quality and Usability of Spoken Dialogue Services*. Speech Communication, 50:730–744, 2008.
- [13] MONITZ, H., A. I. MATA, I. TRANCOSO und M. C. VIANA: *How can you use disfluencies and still sound as a good speaker?*. In: Proc. INTERSPEECH, S. 1678, 2008.
- [14] PUCKETTE, M.: *The Theory and Technique of Electronic Music*. <http://puredata.info/>, 2007.
- [15] RAAKE, A., C. SCHLEGEL, K. HOELDTKE, M. GEIER und J. AHRENS: *Listening and conversational quality of spatial audio conferencing*. In: Proc. 40th Conference of Audio Engineering Society (AES), Bd. 3, 2010.
- [16] SCHUBERT, A. und W. SENDLMEIER: *Was kennzeichnet gute Nachrichtensprecher im Hörfunk? Eine perzeptive und akustische Analyse von Stimme und Sprechweise*. In: SENDLMEIER, W. (Hrsg.): *Sprechwirkung – Sprechstile in Funk und Fernsehen*, Nr. 3 in *Mündliche Kommunikation*, S. 13–70. Logos, Berlin, 2005.
- [17] SCHULZ v. THUN, F.: *Miteinander Reden: Allgemeine Psychologie der Kommunikation*. Bechtermünz, Augsburg, 2000.
- [18] STRANGERT, E. und J. GUSTAFSON: *What makes a good speaker? Subjective ratings, acoustic measurements and perceptual evaluation*. In: Proc. INTERSPEECH, S. 1688–1691, 2008.
- [19] TROUVAIN, J., S. SCHMIDT, M. SCHRÖDER, M. SCHMITZ und W. BARRY: *Modelling personality features by changing prosody in synthetic speech*. In: Proc. Speech Prosody, 2006.
- [20] WALKER, M., C. KAMM und D. LITMAN: *Towards developing general models of usability with PARADISE*. Natural Language Engineering, 6:363–377, 2000.
- [21] WEISS, B. und F. BURKHARDT: *Voice Attributes Affecting Likability Perception*. In: Proc. INTERSPEECH, S. 1934–1937, 2010.
- [22] WEISS, B. und F. BURKHARDT: *Is ‘not bad’ good enough? Aspects of unknown voices’ likability*. In: Proc. Interspeech, S. accepted, Portland, USA, 2012.
- [23] ZUCKERMANN, M., H. HODGINS und K. MIYAKE: *The vocal attractiveness stereotype: Replication and elaboration*. Journal of Nonverbal Behaviour, 14:97–112, 1990.