

AKUSTISCHE KORRELATE VON SYMPATHIEURTEILEN BEI HÖRERN GLEICHEN GESCHLECHTS

Benjamin Weiss

*Quality and Usability Lab, TU Berlin
benjamin.weiss@tu-berlin.de*

Abstract: Konsistente Sympathiebewertungen wurden auf Basis eines Satzes von gleichgeschlechtlichen Hörern abgegeben. Diese korrelieren mit prosodischen Parametern, die aus den Stimuli extrahiert wurden. Während die mittlere Grundfrequenz bei männlichen Sprechern keine Signifikanz erreicht, bestätigt sich der Zusammenhang zwischen positiver Bewertung und einem Tempomaß für weibliche Sprecher. Für spektrale Maße wurde erstmals das Konzept des Sprecherformanten für männliche Sprecher berücksichtigt, dessen Operationalisierung mit der Sympathie korreliert. Für Sprecherinnen zeigt sich ein Ergebnis für die spektrale Schiefe im betonten Vokal.

1 Einleitung

Nach Scherers modifizierter Theorie von Brunswicks Linsenmodell lassen sich distale von proximalen Merkmalen unterscheiden [17]. Für die vokale Kommunikation stellen distale Merkmale externalisierte Signale einer Sprechereigenschaft dar (beispielsweise eines Persönlichkeitseigenschaft oder Emotion), die üblicherweise aus dem Sprachsignal extrahiert werden (z.B. die mittlere Grundfrequenz). Bei einem proximalen Merkmal handelt es sich dagegen um einen Sinneseindruck, bspw. die wahrgenommene Tonhöhe, aufgrund dessen ein Hörer zu einem Urteil oder zu einer Kategorisierung kommt. Solche Urteile, die auf Basis von Höreindrücken gefällt werden, sind Sprecherzuschreibungen, etwa zum Alter oder zur Persönlichkeit, deren Validität erst einmal zu überprüfen ist. Aber auch Einstellungen des Hörers gegenüber dem Sprecher leiten sich aus solchen Urteilen ab.

Ziel dieses Beitrages ist die Erforschung von akustischen distalen Merkmalen, die mit einer positiven oder negativen Einstellung eines Hörers gegenüber unbekanntem Sprechern korrelieren.

2 Forschungsstand

Sympathie oder Aversion stellt die Einstellung gegenüber einem Sprecher dar. Diese Valenzdimension hat durchaus Relevanz für das eigene Verhalten (Annäherung oder Vermeidung) und kann damit bereits einen ersten Schritt der Beziehungsentwicklung darstellen [11].

Diese hörerseitige Einschätzung von Sympathie oder Aversion gegenüber unbekanntem Sprechern wird selten deutlich von proximalen Merkmalen oder Sprecherattributionen unterschieden: So werden proximale Perzepte mittels Fragebögen erhoben, in denen angegeben wird, wie freundlich, angenehm und sympathisch ein Sprecher klingt, obwohl „freundlich“ ein Sprecherattribut, „angenehm“ eine Eindruckswirkung auf den Hörer und „sympathisch“ eine Einstellung gegenüber dem Sprecher darstellt.

Es kann jedoch durchaus valide sein, zu diesem frühen Beziehungsstadium Freundlichkeit, Annehmlichkeit und Sympathie erst einmal nicht zu differenzieren. Systematische Untersuchungen zum Verhältnis von Freundlichkeitszuschreibung, Sympathiebewertung und angenehmer Stimme sind nicht bekannt. Jedoch sollen hier drei grobe Forschungsrichtungen

bezüglich der Valenz unterschieden werden: Persönlichkeitsforschung mit dem 2-Faktorenmodell [1], mit dem 5-Faktorenmodell [12], sowie die Einstellungsforschung [9].

1. Bei dem 2-Faktorenmodell werden je nach Autor die Aspekte Wärme/Wohlwollen/Gemeinschaft von (Handlungs-) Kompetenz/Status unterschieden. Wohlwollen wird üblicherweise mit Fragebogenitems erhoben, die „sympathisch“, „angenehm“ und „freundlich“ beinhalten [4, 14]. Jedoch liegen hierzu auch widersprechende Ergebnisse vor. So lädt bei Weirich [24] „sympathisch“ auf beiden Faktoren ähnlich stark, und in der 2-Faktorenlösung von Street & Brady [20] wird „likeable“, also „sympathisch“, weder in den Faktor Kompetenz, noch in den zweiten Faktor, dort „social attractiveness“ genannt, eingeschlossen. Es zeigen jedoch auch Unterschiede zwischen zwei Stichproben, die auf vergleichbare Weise erhoben und analysiert wurden: Für eine Stichprobe korrelieren Wohlwollen und Kompetenz, für die andere lediglich Wohlwollen mit expliziten „Liking“ Bewertungen [29].
2. Für das 5-Faktorenmodell lässt sich nicht direkt Sympathie ableiten, da es sich hierbei ja um eine Hörereinstellung handelt. Es könnte jedoch vermutet werden, dass eine verträgliche Persönlichkeit, also eine mitfühlenden, kooperative Person, eher sympathisch gefunden wird. Die Ergebnisse der Einstellungsforschung im dritten Punkt zeigen jedoch, dass nicht direkt aus dem 5-Faktorenmodell von einer Persönlichkeitsausprägung auf die ihr gerichtete Sympathie geschlossen werden kann.
3. In der Einstellungsforschung werden interpersonelle Bewertungen vornehmlich für bekannte Personen, also für bereits etablierte Beziehungen, erhoben. Hier wird explizite Sympathie („liking“, „social attractiveness“) sowohl direkt über eine einzelne Skala, als auch anhand von Fragebögen erhoben. In diesem Rahmen wird u.a. untersucht, welche Persönlichkeitsausprägungen von Bewertetem und Urteiler mit der Sympathie korrelieren, wobei hierfür wiederum die 2- oder 5-Faktorenmodelle genutzt werden können. Ergebnisse für solche Beziehungen zeigen zum Beispiel einen positiven Effekt für mitfühlende, aber auch extrovertierte und emotional stabile (Neurotizismus) Personen [23], sowie ähnlicher Persönlichkeiten [22], oder auch physischer Attraktivität. Untersuchungen dieser Einstellungsforschung für unbekannte Sprecher sind jedoch nicht verbreitet, obwohl Attraktivität und einige der 5 Faktoren anhand von rein auditiven Stimuli zugeschrieben werden [18, 30]).

Da lediglich zum 2-Faktorenmodell Ergebnisse für den ersten akustischen Eindruck vorliegen, sollen hier solche Studien für unbekannte Sprecher dargestellt werden, die „sympathisch“ oder „angenehm“ als Items in ihren Fragebögen/Skalen verwendet haben. „Angenehm“ wird mit berücksichtigt, da es in zahlreichen Studien zum 2-Faktorenmodell mit sympathisch auf einem Faktor lädt [4].

Der Wärmefaktor ist höher für weibliche Stimmen, die bewusst als behaucht produziert wurden [14]. Für männliche Sprecher ergaben sich kaum Unterschiede, jedoch wurde die Knarrstimme am negativsten beurteilt.

Zahlreiche Untersuchungen belegen unterschiedliche Systematiken für die Beziehung zwischen Sprechtempo und den beiden Faktoren Wärme und Kompetenz: Während die Attribuierung von Kompetenz mit steigendem Tempo korreliert, zeigt Wärme hingegen einen Idealpunkt-Zusammenhang mit moderatem Tempo [2,4,5,19,20]. Für das Sprechtempo konnte auch ein Effekt des Urteilers nachgewiesen werden, da Hörer solche Sprecher bevorzugen, die ein ähnliches Sprechtempo aufwiesen [7,21].

Zumindest für männliche Sprecher ist auch ein negativer Zusammenhang zwischen Tonhöhe (als mittlere Grundfrequenz, F_0 , gemessen) und Wärme- [4,15], Annehmlichkeits- [6,10] und Sympathiebewertungen (auch für weibliche Sprecher) bekannt [25,28]. Außerdem wird mehr

F₀-Variabilität positiv beurteilt [4,6]. Für die Bewertungen von „Liking“, also Sympathie, kann Gravano [8] den positiven Effekt von mittlere F₀ bestätigen, jedoch nicht für F₀-Range, der sich negativ auswirkt. Für einen kleinen Datensatz („Nijmegen“ [3]), der anhand eines umfangreicheren Fragebogen bewertet wurde, fallen die Items „angenehm“ und „schön“ mit dem Faktor „Melodiös“ zusammen. Allerdings korrelieren hierbei nicht F₀-Variabilität, sondern nur F₀-Range mit diesen beiden Valenz-Items. In einem zweiten Datensatz tritt lediglich eine Korrelation mit dem Anteil stimmhafter Abschnitte zur Gesamtdauer auf.

So zeigen diese Arbeiten, dass ein erster akustischer Eindruck zumindest zu vorläufiger Sympathie bzw. Aversion führen kann. Hierbei werden tiefe Männerstimmen, ein ähnliches bzw. mittleres Sprechtempo und höhere F₀-Variabilität positiv bewertet, wobei die Ergebnisse für F₀-Range nicht einheitlich sind. Dabei sind die akustischen Korrelate anscheinend geschlechtsabhängig, da bspw. in eigenen Arbeiten positiv bewertete Sprecherinnen lediglich ein höheres Sprechtempo aufweisen, während die Sprecher mit tieferer Grundfrequenz sympathischer beurteilt wurden [25].

Weitere akustische Parameter, die sich auf das Spektrum beziehen, wie etwa Harmonic-To-Noise-Ratio (HNR) [8, 27], Spectral Tilt [27] oder Spectral Center of Gravity (CoG) [27] zeigen häufig keinen signifikanten Zusammenhang mit Sympathiebewertungen, sind jedoch auch bei Weitem nicht so gut untersucht wie F₀- und Tempomaße. Da es hier aber Inkonsistenzen gibt, bspw. für einen Zusammenhang zwischen Sympathie-Bewertungen und dem spektralen Maß CoG [26], sowie dagegen, [27], sollen hier insbesondere potentielle spektrale Korrelate von Sympathie betrachtet werden.

Einen solcher potentieller Parameter stellt der so genannte „Speaker’s Formant“ [13] dar. In Anlehnung an die bei professionellen klassischen Sängern bekannte lokale Energiemaximum im Long-Term-Average-Spectrum (LTAS) im Bereich 2.5kHz befindet sich der Speaker’s Formant bei Männern etwa 1kHz höher und, nach der eher spärlichen Anzahl von Untersuchungen, nicht so ausgeprägt wie beim Singer’s Formant. Dennoch zeigt sich in der Stärke des Maximums eine Unterscheidbarkeit zwischen professionellen und unprofessionellen Sprechern, deren artikulatorische Ursache jedoch noch diskutiert wird. Der Speaker’s Formant wird in dieser Untersuchung gemeinsam mit anderen spektralen Maßen wie CoG und Spectral Tilt ausgewertet, wobei ein positiver Zusammenhang mit der Stärke dieser Resonanz lediglich für männliche Sprecher erwartet wird.

3 Durchführung

Jeweils 13 weibliche und 13 männliche Sprecher wurden von je 15 Hörern des gleichen Geschlechts auf einer einzelnen 7-stufigen Skala auf Sympathie bewertet. Diese Bewertungen wurden direkt im Anschluss an ein Experiment zur Sprecherunterscheidung für dieselben Sprecher durchgeführt, sodass die Hörer bereits mit diesen Sprechern leicht vertraut waren, und somit längere oder mehrere Stimuli für einen ersten Eindruck nicht mehr nötig waren. Die gesamte Durchführung dauerte etwa 55 Min., die Sympathiebewertung etwa 5 Min.

Im Vorfeld wurden den Teilnehmern jeweils Triples des gleichen Satzes dargeboten, die von unterschiedlichen Sprechern geäußert wurden. Die Aufgabe bestand darin, zwei ähnlich klingende Sprecher zu identifizieren, die sich von einem dritten Stimulus unterscheiden. Die Ähnlichkeit des Paares und der Unterschied zum dritten Stimulus wurden in einem Textfeld anhand von Adjektiven (oder kurzen Stickworten) beschrieben, und eine Präferenz für eines der Beschreibungen angegeben. 52 dieser Triples wurden pro Teilnehmer bearbeitet, jeweils mit einem neuen Satz aus der Phondat 1 Datenbank.

Nach einer kurzen Pause (1–2) Min. wurde ein Persönlichkeitsfragebogen ausgefüllt (BFI-S, 15 Items) und als Abschluss die Sympathiebewertung mit LisTEN durchgeführt [16]. Hierbei

wurde ein 53. Satz („Heute ist schönes Frühlingswetter.“) über Kopfhörer (AKG K-601) dargeboten: Die Reihenfolge wurde dabei automatisch randomisiert. Als Instruktion wurde gebeten, auf der 7-stufigen Skala die Sympathie gegenüber dem Sprecher bzw. der Sprecherin abzugeben, und das dieser neue Satz als Grundlage dazu diene.

Für die Phondat 1 Sprecher sind keine weitergehenden Sprecherinformationen vorhanden, aber die männlichen Teilnehmer waren im Alter von 22–46 Jahren ($M=29,8$, $SD=7,66$) und die weiblichen waren 21–36 Jahre alt ($M=27,3$, $SD=4,25$). Kein Teilnehmer kam aus einem professionellen Bereich der Phonetik oder Psychologie. Alle wurden für Ihre Teilnahme entschädigt.

4 Messungen

Akustische Messungen des Stimulussatzes wurden mit Praat durchgeführt. Insgesamt wurden sechs Parameter erhoben, um prosodische Merkmale als potentielle Korrelate der Sympathieurteile zu erheben.

1. *Mittlere F₀* über den gesamten Stimulus; empfohlene Standardeinstellungen getrennt für Männer und Frauen.
2. *F₀ Varianz* über den gesamten Stimulus; als Maß der Variabilität der Tonhöhe
3. *Sprechdauer* zur Erfassung des Sprechtempos
4. „Sprecherformant“: Das LTAS wird mit 100Hz Bändern über den gesamten Stimulus berechnet. Die Energie des lokalen Maximums im Bereich 3–4kHz bzw. 4–5kHz wird relativ zur Gesamtenergie bestimmt.
5. Center of Gravity (CoG) für den betonten Vokal /ε/ in Frühlingswetter.
6. Spektrale Schiefe (Skw) für den betonten Vokal /ε/ in Frühlingswetter.

5 Ergebnisse

Die Konsistenz der Urteile ist mit einer Intra-Class Correlation von 0,85 (Männer) und 0,89 (Frauen) zufriedenstellend und vergleichbar mit ähnlichen Untersuchungen.

Es ergeben sich mittlere und starke Spearman Korrelationen zwischen der Rangfolge sympathischer Sprecher mit den akustischen Maßen.

Für die Sprecher ergeben sich nennenswerte Korrelationen für die mittlere Grundfrequenz und den Speaker's Formant (Tabelle 1). Jedoch ist die mittlere F₀ mit $p=.053$ nicht signifikant.

Tabelle 1 – Ergebnisse der Rangkorrelationen.

Parameter	Männer	Frauen
F ₀ (Mittelwert)	rho=-0,55	rho =-0,16
F ₀ (Varianz)	rho=-0,25	rho =-0,18
Dauer	rho= 0,07	rho =-0,61*
Max LTAS[3-4kHz]/Mean	rho= 0,61*	rho = 0,27
CoG	rho= 0,25	rho =-0,30
Skw	rho= 0,03	rho = 0,63*

Eine multiple, lineare Regression schließt nach AIC beide Parameter mit ein (Abb. 1, $R^2=0,47$, $p=0,04$).

Für die Sprecherinnen zeigen sich jedoch andere Zusammenhänge, nämlich Korrelationen mit dem Sprechtempo und der spektralen Schiefe, die beide signifikant sind. Eine multiple Regression beinhaltet nach AIC jedoch lediglich die Dauer, nicht mehr die spektrale Schiefe (Abb. 1, $R^2=0,48$, $p=0,04$).

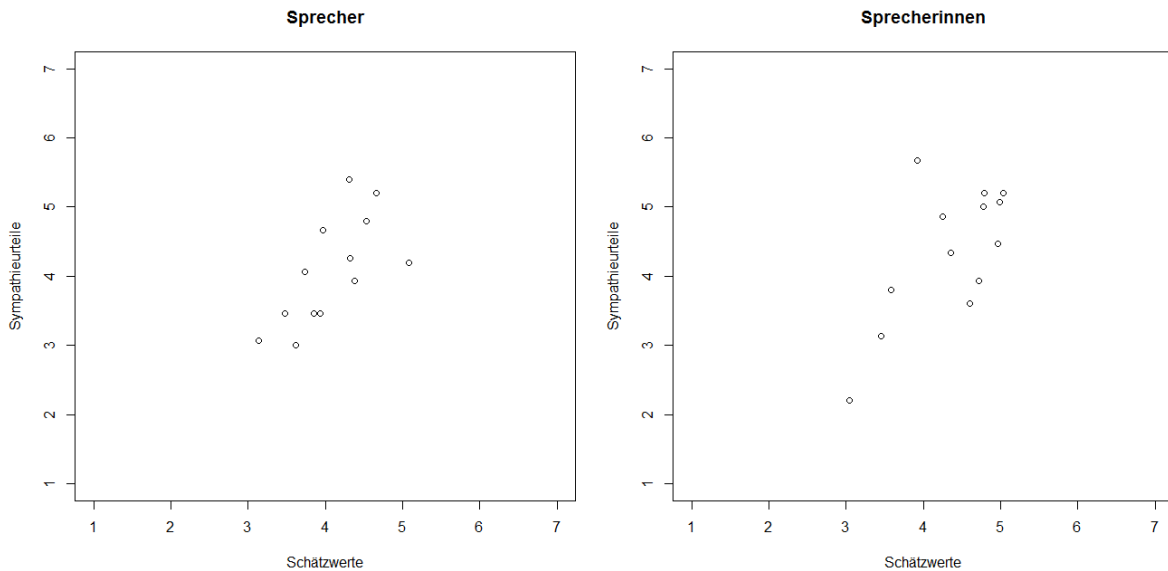


Abbildung 1 – Ergebnisse der linearen Modelle.

6 Diskussion und Fazit

Die beiden Gruppen der je 13 Sprecher/innen wurden trotz der recht kurzen Stimulusdauern konsistent beurteilt. Da dies auch an der Vertrautheit mit den Sprechern durch das vorgelagerte Experiment liegen kann, kann nicht geschlossen werden, dass solch kurze, einzelne Stimuli zu derart konsistenten Resultaten führen. Beispielweise können Sprecher mit unterschiedlichen Sätzen wiederholt dargeboten werden, um reliable Bewertungen zu erhalten [26], während Einzeldarbietungen durchaus zu deutlich geringer Konsistenz führen können ([27], hier jedoch mit niedrigerer Aufnahmequalität).

Für die männlichen und weiblichen Sprecher ergeben sich unterschiedliche signifikante Korrelationen zwischen den Sympathiebewertungen und akustischen Parametern. Der häufig gefundene negative Zusammenhang mit mittlerer F_0 bei Männern erreicht jedoch nicht Signifikanz, was noch einmal die Grenzen einer niedrigen Sprecheranzahl vor Augen führt. Die Relevanz von Sprechtempo bei Sprecherinnen konnte jedoch bestätigt werden, ein Einfluss von F_0 -Variabilität jedoch nicht.

Für spektrale Maße zeigt sich der vermutete positive Effekt von spektraler Energie im Bereich des Sprecherformanten für Männer, sowie ein Effekt der spektralen Schiefe für Sprecherinnen. Eine Bedeutung des spektralen Center of Gravity hingegen wurde nicht bestätigt.

Während das Konzept des Sprecherformanten bereits Grundlagen für eine Interpretation des positiven Zusammenhanges auf die Sympathie beinhaltet (effizientere Stimmanregung, „gute“ Stimmqualität), sind Ergebnisse für andere spektrale Maße nicht nur aufgrund der uneindeutigen empirischen Ergebnisse schwieriger zu interpretieren. Hier sollte mit Hilfe subjektiver Stimmbeschreibungen und einer holistischen Erfassung von Eindruckswirkung

überprüft werden, warum eine entsprechende spektrale Ausprägung zu Sympathie bzw. Aversion führt. Der den hier vorgestellten Bewertungen vorgelagerte Teil des Experimentes wird hierzu hinsichtlich einer Fragebogengestaltung ausgewertet.

Danksagung

Diese Arbeit wurden von der DFG gefördert (WE 5050/1-1).

Literatur

- [1] ABELE, A.E., CUDDY, A.J.C., JUDD, C.M., YZERBYT, V.Y.: Fundamental dimensions of social judgment. editorial to the special issue. *European Journal of Social Psychology* 38 (7), 2008, S. 1063–1065.
- [2] APPLE, W., STREETER, L.A., KRAUSS, R. M.: Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology* 37(5), 1979, S. 715–727.
- [3] BOVES, L.: *The Phonetic Basis of Perceptual Ratings of Running Speech*. Foris, Dordrecht, 1984.
- [4] BROWN, B.L., STRONG, W.J., RENCHER, A.C.: Fifty-four voices from two: the effects of simultaneous manipulations of rate, mean fundamental frequency, and variance of fundamental frequency on ratings of personality from speech. *Journal of the Acoustical Society of America* 55(2), 1974, S. 313–318.
- [5] BROWN, B.L., STRONG, W.J., RENCHER, A.C.: Acoustic determinants of perceptions of personality from speech. *Linguistics* 13 (166), 1975, S. 11–32.
- [6] BRUCKERT, L., LIENARD, J., LACROIX, A., KREUTZER, M., LÉBOUCHER, G.: Women use voice parameter to assess men's characteristics. *Proc. Biological Sciences* 237, 2006, S. 83–89.
- [7] FELDSTEIN, S., DOHM, F., CROWN, C.: Gender and speech rate in the perception of competence and social attractiveness. *Journal of Social Psychology* 141, 2001, S. 785–806.
- [8] GRAVANO, A., LEVITAN, R., WILLSON, L., BENUŠ, Š., HIRSCHBERG, J., NENKOVA, A.: Acoustic and prosodic correlates of social behavior. In: *Proc. Interspeech*, 2011, S. 97–100.
- [9] HERKNER, W.: *Lehrbuch Sozialpsychologie*. Huber, Bern, 1991.
- [10] KETZMERICK, B.: *Zur auditiven und apparativen Charakterisierung von Stimmen*. Studentexte zur Sprachkommunikation. TUDpress, Dresden, 2007.
- [11] LEVINGER, G., SNOEK, J.D.: *Attraction in relationship: A new look at interpersonal attraction*. General Learning Press, Morristown, N.J., 1972.
- [12] MCCRAE R.R., TERRACCIANO A.: Personality profiles of cultures: aggregate personality traits. *Journal of Personality and Social Psychology* 89(3), S. 407–25.
- [13] NAWKA, T., ANDERS, L.C., CEBULLA, M., ZURAKOWSKI, D.: The speaker's formant in male voices. *Journal of Voice* 11(4), 1997, S. 422–428.
- [14] PITTAM, J.: Listeners' Evaluations of Voice Quality in Australian English Speakers. *Language and Speech* 30, 1987, S. 99–113.
- [15] RAY, G.B.: Vocally cued personality proto-types: An implicit personality theory approach. *Communication Monographs*, 53, 1986, S. 266–276.
- [16] SCHÄFER, M., SCHNELLING, C., GEISER, B., AND VARY, P.: A Listening Test Environment for Subjective Assessment of Speech and Audio Signal Processing Algorithms. *Konferenz Elektronische Sprachsignalverarbeitung (ESSV)*, TUDpress, Dresden, 2011, S. 237–244.
- [17] SCHERER, K.R.: Personality inference from voice quality: The loud voice of extroversion. *European Journal of Psychology* 8, 1978, S. 467–487.
- [18] SCHERER, K.R.: Personality markers in speech. In: SCHERER, K.R., GILES, H. (Hrsg.), *Social Markers in Speech*. Cambridge University Press, Cambridge, 1979, S. 147–209.

- [19] SMITH, B.L., BROWN, B.L., STRONG, W.J., RENCHER, A.C.: Effects of speech rate on personality perception. *Language and Speech*, 18, 1975, S. 145–253.
- [20] STREET, R.L., JR., BRADY, R.M.: Speech rate acceptance ranges as a function of evaluative domain, listener speech rate, and communicative Context. *Communication Monographs*, 49, 1982, S. 290–308.
- [21] STREET, R.L.J., BRADY, R.M., PUTNAM, W.B.: The influence of speech rate stereotypes and rate similarity on listeners' evaluation of speakers. *Journal of Language and Social Psychology* 2(1), 1983, 37–56.
- [22] TENNEY, E., TURKHEIMER, E., OLTMANN, T.: Being liked is more than having a good personality: The role of matching. *Journal of Research in Personality* 43, 2009, S. 578–585.
- [23] VAN DER LINDEN, D., SCHOLTE, R.H., CILLESSEN, A.H., TE NIJENHUIS, J., SEGERS, E.: Classroom ratings of likeability and popularity are related to the Big Five and the general factor of personality. *Journal of Research in Personality* 44, 2010, S. 669–672.
- [24] WEIRICH, M.: *Die attraktive Stimme: Vocal Stereotypes. Eine phonetische Analyse anhand akustischer und auditiver Parameter*. Saarbrücken: Verlag Dr. Müller, 2010.
- [25] WEISS, B.: Prosodische Elemente vokaler Sympathie. In: *24th Konferenz Elektronische Sprachsignalverarbeitung (ESSV)*, Hrsg. von P. WAGNER. Bd. 65. Studentexte zur Sprachkommunikation. Dresden: TUDpress, 2013, S. 212–217.
- [26] WEISS, B., BURKHARDT, F.: Voice Attributes Affecting Likability Perception. *Proc. Interspeech*, 2010, S. 1934–1937.
- [27] WEISS, B., BURKHARDT, F.: Is 'not bad' good enough? Aspects of unknown voices' likability. In: *Proc. Interspeech*, 2012.
- [28] WEISS, B., MÖLLER, S.: Wahrnehmungsdimensionen von Stimme und Sprechweise. In: *Proc. Elektronische Sprachsignalverarbeitung (ESSV)*, Aachen, KRÖGER, B.J. & BIRKHOLZ, P. (Eds.), Studentexte zur Sprachkommunikation 61, Dresden: TUDpress, 2011, S. 261–268.
- [29] WORTMAN, J., WOOD, D.: The personality traits of liked people. *Journal of Research in Personality* 45, 2011, S. 519–528.
- [30] ZUCKERMANN, M., DRIVER, R.: What sounds beautiful is good: the vocal attractiveness stereotype. *Journal of Nonverbal Behaviour* 13, 1989, S. 67–82.