



Case Report: Women, Be Aware that Your Vocal Charisma can Dwindle in Remote Meetings

Ingo Siegert^{1*} and Oliver Niebuhr²

¹Mobile Dialog Systems, Institute for Information Technology and Communication, Otto von Guericke University Magdeburg, Magdeburg, Germany, ²Centre for Industrial Electronics, University of Southern Denmark, Sønderborg, Denmark

Remote meetings *via* Zoom, Skype, or Teams limit the range and richness of nonverbal communication signals. Not just because of the typically sub-optimal light, posture, and gaze conditions, but also because of the reduced speaker visibility. Consequently, the speaker's voice becomes immensely important, especially when it comes to being persuasive and conveying charismatic attributes. However, to offer a reliable service and limit the transmission bandwidth, remote meeting tools heavily rely on signal compression. It has never been analyzed how this compression affects a speaker's persuasive and overall charismatic impact. Our study addresses this gap for the audio signal. A perception experiment was carried out in which listeners rated short stimulus utterances with systematically varied compression rates and techniques. The scalar ratings concerned a set of charismatic speaker attributes. Results show that the applied audio compression significantly influences the assessment of a speaker's charismatic impact and that, particularly female speakers seem to be systematically disadvantaged by audio compression rates and techniques. Their charismatic impact decreases over a larger range of different codecs; and this decrease is additionally also more strongly pronounced than for male speakers. We discuss these findings with respect to two possible explanations. The first explanation is signal-based: audio compression codecs could be generally optimized for male speech and, thus, degrade female speech more (particularly in terms of charisma-associated features). Alternatively, the explanation is in the ears of the listeners who are less forgiving of signal degradation when rating female speakers' charisma.

Keywords: prosody, charisma, sex differences, phonetics, distant meetings, degraded speech, audio codec

OPEN ACCESS

Edited by:

Daniel Broudy,
Okinawa Christian University, Japan

Reviewed by:

Barbara Ann Bush,
University of California,
United States

*Correspondence:

Ingo Siegert
ingo.siegert@ovgu.de

Specialty section:

This article was submitted to
Political Communication
and Society,
a section of the journal
Frontiers in Communication

Received: 29 September 2020

Accepted: 16 December 2020

Published: 25 January 2021

Citation:

Siegert I and Niebuhr O (2021) Case
Report: Women, Be Aware that Your
Vocal Charisma can Dwindle in
Remote Meetings.
Front. Commun. 5:611555.
doi: 10.3389/fcomm.2020.611555

1 INTRODUCTION

According to the definition of charisma by (Michalsky, 2019), the power of charismatic speakers originates from three key abilities: The ability to evoke trust by conveying competence, the ability to evoke inspiration by conveying passion, and the ability to evoke motivation by conveying self-confidence. Thus, in listener ratings of speakers, the concept of charisma is—as in the present study—highly correlated with attributes like trustworthy, persuasive, and decided, as well as likeable,

enthusiastic, stimulating, and visionary, see (Rosenberg and Hirschberg, 2009; Niebuhr and Wrzeczsz, 2019; Brem and Niebuhr, 2020).

Perceived charisma is not only a fascinating research subject. It also gives speakers many practical advantages in their everyday life. Research shows that charismatic speech makes listeners more attentive and, in addition, increases the willingness of, for example, employees or students to learn, act, and work in an effective and committed manner (Antonakis et al., 2011; Lee et al., 2014; Towler et al., 2014). Creativity workshops, which are led by more charismatic moderators, end with a significantly better and qualitatively higher idea output (Bachleitner, 2018) and, thus, contribute to the competitiveness and innovative strength of companies or societies.

The overall impression of a charismatic speaker is the result of a complex interplay of many parameters. They range from age and gender (Jokisch et al., 2018) through clothing (Brem and Niebuhr, 2020) and choice of words (Antonakis et al., 2011) to body language and voice (Scherer et al., 2012; Wörtwein et al., 2015; Caspi et al., 2019). The voice seems to play a particularly important role in this interplay. Experimental studies show that the voice not only allows predicting significantly (with 70–80 percent correctness) which idea presentation in startup contests will receive an investment and which will not. Voice analyses can also indicate how high the investment will be (Schweißfurth et al., 2020). Implemented in robots, voices with more charismatic parameter settings make listeners fill out longer questionnaires, book certain travel destinations, and even take detours by car (Fischer et al., 2019; Niebuhr and Michalsky, 2019). Vocal charisma signals include virtually all aspects of a speaker's voice, but are particularly associated with pitch (i.e., fundamental frequency, f_0) level and variability, speaking rate, and aspects of voice quality in terms of both loudness and spectral-energy distribution measures. Experimental-phonetic research shows that higher parameter levels (i.e., e.g., higher levels of pitch and speaking rate) make speakers sound more charismatically in the ears of listeners. The only except are voice-quality measures of spectral energy distribution. Here, it is often lowers values, i.e., a more balanced or flat spectral-energy distribution, that increase a speaker's charismatic impact on listeners (Rosenberg and Hirschberg, 2009; Niebuhr et al., 2018; Fischer et al., 2019; Schweißfurth et al., 2020). Regarding speaker gender, previous findings suggest that, all else equal, women's voices have a less charismatic impact than men's voices. This applies in particular to persuasion-oriented settings like startup funding contests or similar investment decision-making and business tasks (Brooks et al., 2014; Niebuhr et al., 2018; Niebuhr and Skarnitzl, 2019; Niebuhr and Wrzeczsz, 2019; Brem and Niebuhr, 2020) Since the all-else-equal principle does not usually apply in everyday life, women do not necessarily have a disadvantage. Especially since women can train your vocal parameters faster and use them more flexibly in many situations (Reichel and Beňuš, 2018; Niebuhr et al., 2019), the acoustic cues to perceived speaker charisma are sometimes even more pronounced for women than for men.

Our key point in the present study is that, in remote meetings, voice-related charisma is virtually the only means that remains for charismatic speakers to win over their audience and make their listeners buy into their ideas, actions, or offers. However, at

the same time, the speaker's voice is subject to a more or less strong signal compression, which, in turn, generates more or less audible artifacts. Whether and to what extent these artifacts of signal compression influence the charismatic impact of a speaker's voice has so far hardly been researched, even though remote meetings have been on the rise not only since the Corona pandemic (Maximize Market Research, 2020).

Gallardo (2018) investigated the effects of communication channel bandwidth (narrow-band speech transmission 300–3,400 Hz vs. wide-band speech transmission 50–7,000 Hz) on the perceived warmth and attractiveness of speaker voices. Her results show clear effects of transmission bandwidth on perceived speaker attributes; moreover, she found male and female speakers to be differently affected by these bandwidth effects. Male speakers were overall affected more strongly and consistently. But these effects were restricted to behavioral or interactional attributes that are less relevant for perceived speaker charisma. Specifically, male speakers sounded more childish and less sympathetic under low-quality narrow-band than under high-quality wide-band speech transmission conditions. Female speakers, by contrast, sounded more ugly¹ and submissive and less competent in the low-quality narrow-band speech transmission condition. That is, for them the transmission-bandwidth effects concerned attributes that are more relevant for perceived speaker charisma. This is noteworthy not least because experimental charisma research suggests that women are, all else equal, generally perceived to be less charismatic than men (Brooks et al., 2014; Niebuhr et al., 2018; Niebuhr et al., 2019).

In a follow-up study, Gallardo and Sanchez-Iborra (2019) varied and compared different speech compression codecs, but instead of investigating the effects of these codecs on perceived speaker attributes, they looked at how severely the codecs deteriorated the performance of automatic speaker-classification algorithms. Also, the study by Siegert et al. (2016b) did not directly deal with speaker attributes but investigated the effects of codec-degraded speech on the ability of listeners to identify and distinguish the perceived speakers' emotions. Siegert et al. found that any signal degradation made the speakers' emotions less intelligible for listeners.

Although the variables and findings of Gallardo and colleagues and Siegert and colleagues are only loosely associated with our aim to shed light on the connections between speech signal compression and perceived speaker charisma, they provide a solid empirical basis for putting forward two pairs of hypotheses:

- H (1a): Speech signal compression codecs affect perceived speaker charisma;
- H (1b): This effect is overall unfavorable, i.e., stronger compression means lower charisma;
- H (2a): Male and female speakers' charisma is differently affected by speech signal compression;
- H (2b): Female speakers' perceived charisma is more strongly affected than that of male speakers.

¹The term ugly is a scale-item taken from the utilized semantic differential questionnaire of speaker characteristics [cf. Fernández Gallardo and Weiss (2018)].

We test these 2×2 Hypotheses with German stimuli and listeners in a perception experiment, using scalar ratings. The complex impression of perceived charisma is queried both directly and indirectly *via* closely correlated attributes selected on an empirical basis.

2 STUDY DESIGN

2.1 Stimuli

To have high-quality speech utterances, independent of the speech content, we used the Berlin Database of Emotional Speech (EMO-DB) (Burkhardt et al., 2005). This database comprises German utterances that have a neutral semantic content, but are realized with different emotional prosodies as well as in a neutral matter-of-fact version by 10 professional actors (five female), pseudonymized *via* a speaker-id². It comprises high-quality recordings both in technical (sampling frequency of 16 kHz, stored as uncompressed WAV at 16-bit depth, bit-rate: 256 kBit/s) and acoustic terms (clear sonorous voices, no influencing content).

For our study, we selected a subset of two male (11 and 15) and two female (13 and 14) speakers from whom a constant utterance set was available for all emotions and the neutral version. The selected subset included 26 uncompressed utterances that constituted the first part of our experimental stimuli.

2.2 Audio Codecs

For data transmission, compression is heavily used within modern (mobile/remote) systems. Compression allows reducing the transmission bandwidth while retaining the speech intelligibility (ITU-T, 1996; ITU-T, 2014; Maruschke et al., 2016). Several codecs have been developed to meet various applications with different quality requirements (Siegert et al., 2016a). More details about the degradation of acoustic characteristics under compressed speech can be found in (Byrne and Foulkes, 2004; Lee et al., 2006; Siegert et al., 2016b).

For this study, we selected four wideband/fullband codecs aiming at specific application scenarios. **Adaptive Multi-Rate Wideband (AMR-WB)** is a high-quality audio compression format mainly used in mobile communications (ITU-T, 2003), also known as “HD Voice” and Voice over LTE (VoLTE). We chose a bitrate of 12.65 kBit/s, which is intended for pure speech signals (ITU-T, 2003).

MPEG-1/MPEG-2 Audio Layer III (MP3) is a popular lossy fullband audio codec (Brandenburg, 1999). It uses perceptual coding for audio compression: certain parts of the original signal, considered to be beyond the auditory resolution ability, are discarded. Besides its usage for music storage, lower bitrates (16 kBit/s) are used to encode audio dramas (Ahern, 2020).

OPUS is an open-source lossy audio codec usable for both speech and music (Valin et al., 2012). It further offers a hybrid mode to improve the speech intelligibility at low bit rates, by

enriching the synthesized signal with characteristics represented by a psychoacoustic model (Valin et al., 2013). The application of the hybrid mode can be controlled by the bitrate, which has to be 34 kBit/s in our case.

SPEEX is an open-source fullband speech codec for internet applications requiring particularly low bit rates (Xiph.Org Foundation, 2014). It is also used as a speech codec in common voice assistant platforms (Caviglione, 2015). The encoding is controlled by a quality parameter that ranges from 0 (worst) to 10 (best). In our study, we used 0 (i.e., 3.95 kBit/s).

All 26 uncompressed stimuli have been compressed employing each of the four presented codecs at the specified bit rate (**AMR-WB**: 12.65 kBit/s, **MP3**: 16 kBit/s, **OPUS**: 34 kBit/s, **SPEEX**: 3.95 kBit/s). This resulted in 104 compressed stimuli. The total number of stimuli in our experiment was hence $26 + 104 = 130$ stimuli.

2.3 Procedure

The perception test was conducted *via* an online survey tool [SoSci Survey Version 3.2.03-i, Leiner (2019)]. The samples were presented in a pseudo-randomized order to avoid that similar samples (same speaker and/or encoding quality) directly follow each other.

After a short audio test and an introduction of the variables, the participants (aka labellers) were asked to listen to all stimuli and rate the respective speaker’s performance on five-point Likert scales (ranging from one “not at all” to five “very strong”), see **Figure 2**. As previous studies showed that speaker charisma is a fairly complex concept and not always easy to apply by listeners in rating tasks, we additionally included nine other speaker attributes that are closely related to speaker charisma and derived from the studies of Rosenberg and Hirschberg (2009); Weninger et al. (2011) as well as from the charisma model of Michalsky (2019). Many of these additional attributes have already been successfully applied in charisma rating tasks of previous studies [e.g., Niebuhr and Wrzeczsz (2019); Niebuhr (2020)].

At the end of the rating task, socio-demographic information (age, sex, mother tongue, BFI-S16) was collected from the participants. Including this additional information, the entire perception experiment took about 1 h.

2.4 Participants

Overall, 21 participants took part in the perception experiment (12 female and nine male, all between 19 and 43 years old, mean 24.76 years). All of them were fluent speakers of German with no reported speech-production or -perception disorders. Furthermore the expression of the personality dimensions was comparable among the participants, thus an influence of the personality on the rating could be excluded.

3 EVALUATION AND RESULTS

In this study, we only analyzed the assessment of the neutral stimuli in various codec qualities. The different emotional versions were excluded. We were interested in how women’s and men’s power of persuasion would be affected by perceived acoustic qualities resulting from different compression codecs.

²More details about this database can be found at: <http://emodb.bilderbar.info/index-1024.html>.

TABLE 1 | Average evaluation of men and women's charisma regarding different compressions denoted as percentual loss compared to high-quality uncompressed speech (WAV).

| | WAV | AMRWB (%) | OPUS (%) | MP3 (%) | SPEEX (%) |
|-------|-------------|-----------|----------|---------|-----------|
| Men | 2.63 (100%) | 97 | 109 | 97 | 71* |
| Women | 3.36 (100%) | 95 | 77* | 86* | 62* |

*Significant changes against WAV.

Therefore, we focus here mainly on comparing the results between the male and female speakers.

For the inferential statistics, we used a two-way repeated-measures ANOVA with the fixed factors Compression (four codecs plus one uncompressed baseline) and Speaker Gender (m/f). Speaker Rating (i.e., the scale values 1–5) served as the dependent variable. Note that, because, as expected, all nine additional speaker attributes were significantly correlated (in terms of Pearson PM r) with each other and with the key attribute of perceived speaker charisma at $r(82) > 0.6$ ($p < 0.001$), cf. also (Rosenberg and Hirschberg, 2009; Niebuhr and Wrzeczsz, 2019) for similar correlation results, we pooled the ratings across all 10 scales and treated them as one coherent rating of perceived speaker charisma. Participant was included in the RM-ANOVA as a random factor.

The RM-ANOVA returned a significant main effect of Compression [$F(4,3356) = 135.7, p < 0.0001, \eta_p^2 = 0.87$], a significant main effect of Speaker Gender [$F(1,3356) = 86.3, p < 0.0001, \eta_p^2 = 0.65$] as well as a significant interaction between the two fixed factors Compression \times Speaker Gender [$F(4,3356) = 58.9, p < 0.0001, \eta_p^2 = 0.44$].

As is summarized in **Table 1**, almost all codecs caused a decrease in perceived speaker charisma as compared to the uncompressed baseline condition (WAV), thus resulting in the significant main effect of Compression. Moreover, we can see in the WAV column of **Table 1** that the overall magnitude of perceived speaker charisma differed between men and women, with the women's stimuli received higher charisma ratings than the men's stimuli. This result underlies the main effect of Speaker Gender. This finding is probably caused by the fact that our stimuli were realized by professional actors. Previous studies show that women benefit quicker and overall more from voice training than men (Niebuhr et al., 2019) and that listeners pay more attention to other features than voice when rating men's than women's speeches (Sellnow and Treinen, 2004; Niebuhr and Wrzeczsz, 2019).

Finally, **Table 1** shows that the factor Compression in the form of its four different codec conditions affects the charisma ratings of men and women to different degrees. The two women suffered more from speech compression than the two men. More specifically, speech compression decreased the men's perceived charisma on average by only 6.5% across all compression conditions, lowering it from 100 to 93.5%. Women's perceived charisma, in contrast, was lowered three times as much across all four compression conditions, i.e., by on average 20%, from 100 to 80%. The RM-ANOVA reflects this fact in the significant interaction between Compression and Speaker Gender. Note that, for OPUS (a frequently used web codec), women's charisma is even absolutely "worse" than for men, as men are rated more charismatic under OPUS compression than in the WAV baseline condition. We suppose that this outcome was caused by the underlying codec. It has been already shown for the hybrid operation mode of OPUS that certain emotions can be recognized better (Siegert et al., 2017), which is in some sense comparable to gender differences in the pronunciation of charisma. Furthermore, for men, only the differences between SPEEX and all other codecs are significant (according to Wilcoxon-Wilcox *post-hoc* tests with Bonferroni correction of alpha-error levels), whereas for women, also the comparisons of WAV vs. MP3 and WAV vs. OPUS came out significantly.

In order to study the interaction Compression*Speaker Gender at greater detail, we broke up our integrated charisma rating again and investigated, again based on Wilcoxon-Wilcox pairwise comparisons tests, how many statistically significant rating decreases we find per attribute for each male and female speaker. The theoretical maximum number of significant rating decreases was 100 ($5 \times 2 = 10$ codec versions \times 10 attributes). We found that the total number of statistically significant codec-induced decreases of charisma or charisma-related ratings differed remarkably between men and women. While for the men only 20% of all comparisons showed a significant decrease (at $p < 0.01$), this value rose to almost 30% in the case of the women (at $p < 0.01$). To illustrate this gender-specific difference, **Figure 1** depicts the decreased ratings caused by SPEEX or OPUS for the charisma-related attributes *trustworthy*, *persuasive*, and *likeable* of men and women. On average, the rating decrease caused the codecs amounted to 1.01 scale points for men, as against 1.49 scale points for women. In addition, **Figure 1** exemplifies our finding that the greater decrease of women's charisma under speech compression was not tied to specific attributes related, e.g., to competence, power or attractiveness.

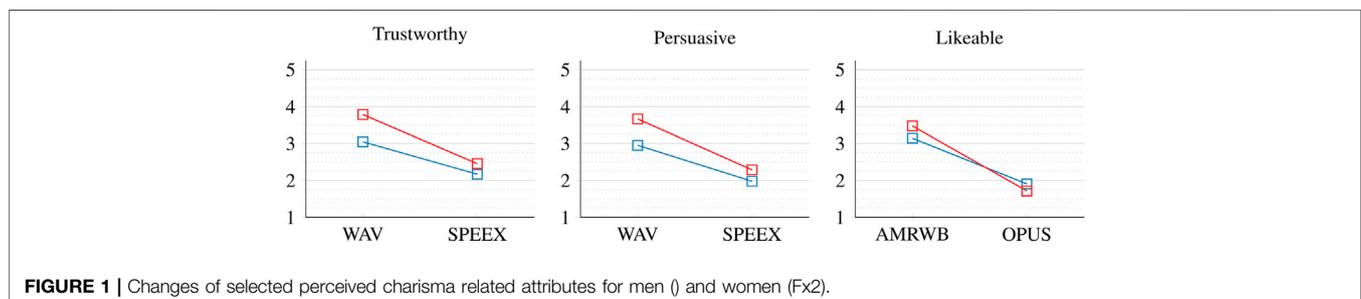


FIGURE 1 | Changes of selected perceived charisma related attributes for men (□) and women (■).

2. Wie wirkt der Sprecher/die Sprecherin auf Sie?
Bitte bewerten Sie dazu folgende Aspekte:

| | überhaupt nicht | wenig | mittel | stark | sehr stark |
|---------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Visionär | <input type="radio"/> |
| Inspirierend | <input type="radio"/> |
| Selbstbewusst | <input type="radio"/> |
| Sympathisch | <input type="radio"/> |
| Attraktiv | <input type="radio"/> |
| Überzeugend | <input type="radio"/> |
| Charismatisch | <input type="radio"/> |
| Glaubwürdig | <input type="radio"/> |
| Entusiastisch | <input type="radio"/> |
| Entschlossen | <input type="radio"/> |

Weiter

FIGURE 2 | Screenshot of utilized annotation questionnaire.

Rather, it occurred across the board and affected all speaker attributes to similar degrees.

4 DISCUSSION AND CONCLUSION

Two pairs of Hypotheses were addressed here, based on previous findings of Siegert and colleagues and Gallardo and colleagues. The results of our perception experiment provide supporting evidence for all four Hypotheses. On this basis, we can draw the following conclusions:

- H (1a): Yes, speech signal compression codecs affect perceived speaker charisma;
- H (1b): Yes, this effect is overall unfavorable, i.e., stronger compression means lower charisma;

- H (2a): Yes, male and female speakers' charisma is differently affected by speech signal compression;
- H (2b): Yes, female speakers' perceived charisma is more strongly affected than that of male speakers.

Note with respect to the latter two conclusions that (Gallardo, 2018) suggested that the stronger negative influence of signal compression on women's voice could be restricted to a few charisma-related attributes like competence and self-confidence (or its counterpart submissiveness). However, this idea is not consistent with our results. They show that the harmful effect of speech compression on women's vocal charisma occurred across the board and concerned all tested charisma-related attributes to a similar degree. Thus, in terms of their relevance for everyday (professional) life, using communication tools with speech compression seems to pose an even more risk for women than previously assumed.

While it is worth pursuing this assumption further in experiments with behavioral or decision-making tasks, the way our data converges with previous evidence leaves hardly any doubt that men and women are not similarly affected by signal compression. This can be either because the compression algorithms treat female speech acoustics worse than male speech acoustics, or because listeners are less forgiving of compression artifacts when they concern female speakers. What argues in favor of the latter is the study of Niebuhr and Wrzeczysz (2019). Its results suggest that the charismatic impact of women more strongly depends on voice-related features than for men. Given that, it is plausible that a degradation of these voice-related features through signal compression particularly lowers the women's charismatic impact. We are currently designing follow-up experiments that take up these more specific questions.

However, irrespective of what the underlying mechanism(s) of the effect may be, the practical implication of our findings is that for women to reach the full charismatic potential it seems crucial to avoid speech-interaction situations with a high risk of signal compression. Even under codecs like Opus or MP3 that are very widespread and not too severe in terms of compression rate, the women's charismatic impact decreases already by about 20% compared to an uncompressed speech-interaction situation (which is at least 10% more than for men). So, for example, women might want to take special care to find a strong network connection before they start their remote meeting and/or rather switch off the video streaming even if it would reduce the bit rate of the audio stream. Note that we make these recommendations based on the neutral matter-of-fact stimuli only, but preliminary analyses of the rating data for the emotional stimuli suggest that the gender-specific compression effect also exists here, probably even more pronounced.

Finally, have to address the women's higher absolute charisma-related ratings compared to the men (Table 1). This difference does not contradict previous research showing that women's speeches are, all else equal, less charismatic than men's speeches. The *ceteris-paribus* principle is just not applicable here as we compared natural stimuli of men and women. That is, the women's higher absolute charisma-related ratings only mean that we by chance selected two women for our 2 + 2 speaker sample,

who spoke at a higher level of vocal charisma than the two men we selected, especially, as the speakers were trained actors. What matters here are the relative differences and changes in charisma-related ratings, and we have started to understand how they look like and how we can potentially overcome them.

We are aware that this study has some limitations, especially the small number of samples as well as the use of acted speech samples. However, our aim was to draw first conclusions and generate valid hypotheses for a clearly defined larger-scale study. Using a small pilot stimulus set ensured a constant, high focus on the task among the raters; and using acted stimuli ensured that we could start from a high charisma level in the uncompressed baseline condition (WAV) so that effects of compression can clearly manifest themselves. In future studies we will include non-acted speech samples (of real video/phone calls) as well as all emotional stimuli that, in combination, are meant to cover the whole spectrum of human feelings in real situations of remote (digital) communication. Based on that, we get a more realistic and differentiated idea about how and to what gender-specific degrees compression in connection with other factors diminishes perceived speaker charisma.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

- Ahern, S. (2020). *Acoustical design of concert halls and theatres: a personal account*. 3rd Edn. Abingdon, United Kingdom: Routledge.
- Antonakis, J., Fenley, M., and Liechti, S. (2011). Can charisma be taught? tests of two interventions. *Acad. Manag. Learn. Educ.* 10, 374–396. doi:10.5465/amle.2010.0012
- Bachsleitner, N. (2018). The role of charismatic moderators for the success of creativity workshops—an experimental study. Master's thesis. Erlangen (Germany): Friedrich-Alexander-Universität Erlangen-Nuremberg.
- Brandenburg, K. (1999). "MP3 and AAC explained," in 17th AES international conference: high-quality audio coding, Florence, Italy, September 2–September 5, 1999.
- Brem, A., and Niebuhr, O. (2020). "Dress to impress? on the interaction of attire with prosody and gender in the perception of speaker charisma," in *Voice attractiveness: concepts, methods, and data*. Editors B. Weiß, J. Trouvain, and J. Ohala (New York, NY: Springer), 301–309.
- Brooks, A. W., Huang, L., Kearney, S. W., and Murray, F. E. (2014). Investors prefer entrepreneurial ventures pitched by attractive men. *Proc. Natl. Acad. Sci. United States.* 111, 4427–4431. doi:10.1073/pnas.1321202111
- Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W., and Weiss, B. (2005). "A database of German emotional speech," in Proc. of the interspeech-2005, Lissabon, Portugal, September 4–September 8, 2005, 1517–1520.
- Byrne, C., and Foulkes, P. (2004). The 'mobile phone effect' on vowel formants. *Int. J. Speech Lang. Law.* 11, 83–102. doi:10.1558/ijsl.v11i1.83
- Caspi, A., Bogler, R., and Tzuman, O. (2019). Judging a book by its cover": the dominance of delivery over content when perceiving charisma. *Group Organ. Manag.* 44, 1067–1098. doi:10.1177/1059601119835982
- Caviglione, L. (2015). A first look at traffic patterns of siri. *Trans. on Emer. Telecom. Technol.* 26, 664–669. doi:10.1002/ett.2697
- Fernández Gallardo, L., and Weiss, B. (2018). "The nautilus speaker characterization corpus: speech recordings and labels of speaker characteristics and voice descriptions (ELRA)," in Proceedings of the eleventh international conference on language resources and evaluation

- (LREC 2018), Miyazaki, Japan, May 7–May 12, 2018 (European Language Resources Association).
- Fischer, K., Niebuhr, O., Jensen, L. C., and Bodenhausen, L. (2019). Speech melody matters—how robots profit from using charismatic speech. *J. Hum.-Robot Interact.* 9 (1), 21. doi:10.1145/3344274
- Gallardo, L. F. (2018). Effects of transmitted speech bandwidth on subjective assessments of speaker characteristics,” in Tenth international conference on quality of multimedia experience (QoMEX) (IEEE), Sardinia, Italy, May 29–May 31, 2018 (IEEE).
- Gallardo, L. F., and Sanchez-Iborra, R. (2019). On the impact of voice encoding and transmission on the predictions of speaker warmth and attractiveness. *ACM Trans. Knowl. Discov. Data.* 13, 4. doi:10.1145/3332146
- ITU-T (1996). *Methods for subjective determination of transmission quality*. Geneva, Switzerland: ITU-T.
- ITU-T (2003). *Wideband coding of speech at around 16 kbit/s using adaptive multi-rate wideband (AMR-WB)*. Geneva, Switzerland: ITU-T.
- ITU-T (2014). Methods for objective and subjective assessment of speech quality (POLQA): perceptual objective listening quality assessment. Doctoral dissertation. Karlskrona (Sweden): Blekinge Institute of Technology.
- Jokisch, O., Iaroshenko, V., Maruschke, M., and Ding, H. (2018). “Influence of age, gender and sample duration on the charisma assessment of German speakers,” in *Studenttexte zur Sprachkommunikation: elektronische Sprachsignalverarbeitung 2018*. Editors A. Berton, U. Haiber, and W. Minker (Saxony, German: TUDpress), 224–231.
- Lee, D.-C., Lu, J.-J., Mao, K.-M., Ling, S.-H., Yeh, M.-C., and Hsieh, C. (2014). Does teachers charisma can really induce students learning interest? *Proc.—Soc. Behav. Sci.* 116, 1143–1148. doi:10.1016/j.sbspro.2014.01.359
- Lee, K.-K., Cho, Y.-H., and Park, K.-S. (2006). “Robust feature extraction for mobile-based speech emotion recognition system,” in Intelligent computing in signal processing and pattern recognition: international conference on intelligent computing, ICIC 2006, Kunming, China, August 16–19, 2006. Editors D.-S. Huang, K. Li, and G. W. Irwin (Berlin, Heidelberg: Springer Berlin Heidelberg), 470–477.
- Leiner, D. J. (2019). Sosci survey version 3.1.06. Available at: <https://www.soscisurvey.de>.
- Maruschke, M., Jokisch, O., Meszaros, M., Trojahn, F., and Hoffmann, M. (2016). “Quality assessment of two fullband audio codecs supporting real-time communication,” in International conference on speech and computer, Budapest, Hungary, August 23–27, 2016, 571–579.
- Maximize Market Research (2020). Report No.: 63052. Global video conferencing market—industry analysis and forecast (2020–2027)—by component, deployment model, organization size, end user, and region. Available at: www.aximizemarketresearch.com (Accessed May, 2020).
- Michalsky, J. (2019). *Einfach überzeugend! teil 1—mehr Charisma durch das Verständnis von Sprache und Stimme*. Bamberg, Germany. Virtuelle Hochschule Bayern (VHB).
- Niebuhr, O. (2020). “Space fighters on stage—how the f1 and f2 vowel-space dimensions contribute to perceived speaker charisma,” in *Studenttexte zur sprachkommunikation: elektronische sprachsignalverarbeitung*. Editors R. Böck, I. Siegert, and A. Wendemuth (Saxony, German: TUDpress), 265–277.
- Niebuhr, O., and Michalsky, J. (2019). “Computer-generated speaker charisma and its effects on human actions in a car-navigation system experiment—or how steve jobs’ tone of voice can take you anywhere,” in Proceedings of computational science and its applications—ICCSA 2019. Editors S. Misra, C. Torre, E. Tarantino, B. O. Apduhan, O. Gervasi, B. Murgante, et al. (Saint Petersburg: Springer-Verlag).
- Niebuhr, O., Skarnitzl, R., and Tylečková, L. (2018). “The acoustic fingerprint of a charismatic voice—initial evidence from correlations between long-term spectral features and listener ratings,” in Proc. 9th international conference on speech prosody, Poznan, Poland, June 13–16, 2018, 359–363.
- Niebuhr, O., and Skarnitzl, R. (2019). “Measuring a speaker’s acoustic correlates of pitch - but which? A contrastive analysis for perceived speaker charisma,” in Proceedings 19th International Congress of Phonetic Sciences, Melbourne, Australia, 1774–1778.
- Niebuhr, O., Tegtmeier, S., and Schweisfurth, T. (2019). Female speakers benefit more than male speakers from prosodic charisma training—a before-after analysis of 12-weeks and 4-h courses. *Front. Commun.* 4, 12. doi:10.3389/fcomm.2019.00012
- Niebuhr, O., and Wrzeczsz, S. (2019). A woman’s gotta do what a woman’s gotta do, and a man’s gotta say what a man’s gotta say—sex-specific differences in the production and perception of persuasive power. Proc. 16th international pragmatics conference. Hong Kong, China, October 2018.
- Reichel, U. D., and Beňuš, S. (2018). Entrainment profiles: comparison by gender, role, and feature set. *Speech Commun.* 100, 46–57. doi:10.1016/j.specom.2018.04.009
- Rosenberg, A., and Hirschberg, J. (2009). Charisma perception from text and speech. *Speech Commun.* 51, 640–655. doi:10.1016/j.specom.2008.11.001
- Scherer, S., Layher, G., Kane, J., Neumann, H., and Campbell, N. (2012). “An audiovisual political speech analysis incorporating eye-tracking and perception data,” in Proc. of the eighth international conference on language resources and evaluation (LREC’12), Istanbul, Turkey, May 23–25, 2012, 1114–1120.
- Schweisfurth, T., Tegtmeier, S., and Niebuhr, O. (2020). “How money talks: speech melody and venture evaluation,” in Proc. 34th conference on research in entrepreneurship and small business, Naples, Italy (Sonderborg, Denmark: Springer-Verlag), 1–10.
- Sellnow, D. D., and Treinen, K. P. (2004). The role of gender in perceived speaker competence: an analysis of student peer critiques. *Commun. Educ.* 53, 286–296. doi:10.1080/0363452042000265215
- Siegert, I., Lotz, A. F., Egorov, O., and Wendemuth, A. (2017). *Improving speech-based emotion recognition by using psychoacoustic modeling and analysis-by-synthesis*. Cham, Switzerland: Springer International Publishing, 445.
- Siegert, I., Lotz, A. F., Duong, L., and Wendemuth, A. (2016a). “Measuring the impact of audio compression on the spectral quality of speech data,” in *Elektronische Sprachsignalverarbeitung 2016*. Leipzig, Germany: TUDpress, 229–236.
- Siegert, I., Lotz, A. F., Maruschke, M., Jokisch, O., and Wendemuth, A. (2016b). ITG-Fb. 267: speech communication: 12. ITG-Fachtagung sprachkommunikation Emotion intelligibility within codec-compressed and reduced bandwidth speech. Paderborn, Germany: VDE Verlag, 215–219.
- Towler, A., Arman, G., Quesnell, T., and Hoffman, L. (2014). How charismatic trainers inspire others to learn through positive affectivity. *Comput. Hum. Behav.* 32, 221–228. doi:10.1016/j.chb.2013.12.002
- Valin, J.-M., Maxwell, G., Terriberry, T. B., and Vos, K. (2013). *The opus codec*. New York, NY: AES International Convention.
- Valin, J.-M., Vos, K., and Terriberry, T. (2012). *Definition of the opus audio codec*. RFC 6716. Available at: <http://www.ietf.org/rfc/rfc6716.txt> (Accessed January 7, 2021).
- Weninger, F., Schuller, B., Batliner, A., Steidl, S., and Seppi, D. (2011). Recognition of nonprototypical emotions in reverberated and noisy speech by nonnegative matrix factorization. *EURASIP J. Appl. Signal Process.* 2011, 16. doi:10.1155/2011/838790
- Wörtwein, T., Chollet, M., Schauer, B., Morency, L.-P., Stiefel, R., and Scherer, S. (2015). “Multimodal public speaking performance assessment,” in Proc. of the 2015 ACM on international conference on multimodal interaction, New York, NY, November 9–13, 2015, 43–50.
- Xiph.Org Foundation (2014). *Speex: a free codec for free speech*. Massachusetts, MA: Xiph.Org Foundation.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Siegert and Niebuhr. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.