



What Makes a Speaker Charismatic? Producing and Perceiving Charismatic Speech

Zixiaofan Yang, Jessica Huynh, Riku Tabata, Nishmar Cestero, Tomer Aharoni, Julia Hirschberg

Columbia University, USA

zy2231@columbia.edu, jyh2127@columbia.edu, rct2125@columbia.edu, nishi@cs.columbia.edu,
ta2453@columbia.edu, julia@cs.columbia.edu

Abstract

Charisma is an essential component of spoken language production and has been used for centuries to engage audiences and obtain followers. Understanding charisma in speech is important not only for text-to-speech synthesis but also for broader issues of explaining social events as well as helping speakers to improve their own charismatic speech production. In this paper, we present the first gender-balanced study of charismatic speech, including speakers and raters from diverse backgrounds. We describe how raters define charisma by analyzing its positive or negative relationship with other speaker traits, such as enthusiasm, persuasiveness, boringness, and uncertainty. Using the features extracted from the voice clips, we analyze the acoustic and textual correlates of charisma. We also extend prior work to examine individual differences in the perception and production of charisma in speech. We discuss how a speaker's gender and how a rater's gender, level of education, personality, and own speaking style influence the rater's perception of charismatic speech.

Index Terms: charisma, speech perception, speaking style

1. Introduction

Charisma was defined by Max Weber as “a certain quality of an individual personality, by virtue of which he is set apart from ordinary men and treated as endowed with supernatural, superhuman, or at least specifically exceptional powers or qualities not accessible to the ordinary person” on which basis “the individual concerned is treated as a leader” [1]. While this definition does not specify the particular qualities that make an individual appear charismatic, previous research has shown some agreement on the personal traits that people associate with charisma [2, 3, 4, 5, 6]. Audiovisual analysis of charisma using video recordings [7, 8] has also found that speech is an essential modality of perceived charisma. Moreover, researchers found that using characteristics of charismatic speech in text-to-speech synthesis can make a computer-generated voice more trustworthy [9] and practicing with acoustic feedback can make humans speak more charismatically [10]; this demonstrates the importance of understanding charismatic speech. However, most previous studies on charismatic speech have examined politicians or industry leaders, focusing on male speakers alone, with relatively few raters and little knowledge of raters' demographic or other information. In this work, we present the first gender-balanced study of charismatic speech, including raters' demographic and personality information, together with samples of raters' own speech, using speakers who are neither politicians nor business leaders. The remainder of the paper is organized as follows: Section 2 describes previous work on charismatic speech. We present details of data collection in Section 3. Section 4 presents results of ratings for charisma and

other speaker traits, together with acoustic-prosodic and lexical analysis of charismatic speech. It also explains how raters' demographics, personality, and their own speech influence their speaker ratings. We conclude in Section 5 and present directions for future work.

2. Related work

One of the early studies of charismatic speech, Rosenberg and Hirschberg [2] [3] collected American political speech segments and asked raters to rate the segments on charisma and 26 additional speaker traits. They found that charismatic speakers used longer sentences, more first-person plural and third-person singular pronouns, more repetitions and complex words; acoustic-prosodic correlates of charismatic speech were higher in pitch, faster, and louder, with more variation in intensity. In a later extension, Biadys et al. [4] studied the cross-cultural perception of charismatic speech and identified many features common across cultures. Also examining political speech, Signorello et al. [5] [6] asked raters to rate an Italian politician's speech for charisma and other 67 traits; D'Errico et al. [11] manipulated the pitch and pause length of Italian and French political speech and collected charisma ratings cross-culturally; Cullen et al. [12] crowd-sourced charisma ratings on an Irish politician's speech and built automatic systems to detect charisma. For charisma in business, Weninger et al. [13] rated charismatic speech from 143 male business executives. Several studies [14] [15] compared the speech of Steve Jobs and Mark Zuckerberg, and found that the more charismatic speaker can be characterized as having a higher F0 level, a larger F0 range, a higher level of variability in speech and a clearer pronunciation. However, when the speech is from male lecturers, people rate low F0 range and low speaking rate as more charismatic [16].

Most research on charismatic speech has focused on the speech of politicians and business leaders, and most speakers rated have been male. To investigate possible gender bias in charismatic speech, Novak et al. [17] compared 1 male and 2 female business executives and found that females produced stronger acoustic charisma cues but were still judged to be as charismatic as the single male speaker. Niebuhr et al. [18] found that female speakers start with significantly lower prosodic-charisma scores than male speakers, judged by an automatic scoring system. However, the charisma cues and scoring metrics in both works were taken from previous literature, which might be already biased towards male speech, without fully understanding the characteristics of female charismatic speech.

In this work, we examine ratings of equal numbers of male and female speakers, also identifying the demographic and personality information of crowd-sourced raters. We want to determine whether raters scores male and female speakers differently when the corpus is balanced for gender, and whether male

or female raters are biased in their ratings of speakers of different genders. We also want to obtain a more detailed study of the lexical and acoustic-prosodic factors significantly correlated with charisma ratings for each gender and to compare how raters rated both on a large number of speaker traits positively or negatively correlated with charisma in previous studies.

3. Data collection

To build a gender-balanced charismatic speech corpus, we selected 30 male and 30 female speech clips from Youtube and pilot tested these to balance charismatic, boring, and neutral groups for each gender using multiple lab ratings. We avoided voice clips from celebrities to prevent rating bias resulting from speaker recognition. The clips were chosen from prepared talks, educational lectures, and interviews, and were each 20 seconds long. Since previous research [12] has found that charisma labels provided by crowd-sourced workers are as reliable as on-site annotators, we used Amazon Mechanical Turk (MTurk) to collect ratings for the 60 voice clips from 15-20 raters each.

Our Human Intelligence Tasks (HITs) were designed as follows: First, workers answered demographic questions, including birth gender, gender preference, and level of education, and completed the Ten Item Personality Inventory (TIPI) [19] to measure their Big-Five personality dimensions [20]. Then, each worker was instructed to rate 10 clips on charisma and 17 other traits: boringness, coldness, confidence, eloquence, enthusiasm, extroversion, fluency, intelligence, introversion, liveliness, ordinariness, persuasiveness, reasonableness, sincerity, trustworthiness, uncertainty and weakness. The clips consisted of 5 voices each from male and female speakers, and the 18 total speaker traits were shuffled multiple times to display different random orders. In addition, a textual attention check instructing workers to select a specific rating and an extra clip served as an audio attention check were mixed in with the other questions and clips, to filter out workers attempting to randomly assign ratings without listening to the voices and ensure the quality of the crowd-sourced data. After completing ratings of the clips, workers were asked to record themselves reading the following passage in their natural voice: “My name is Robin, and after years of working for other startups, I’m taking the plunge and developing my own app. The app allows anyone to rent a car by the hour, without having to go through a rental company. They can pick the car up, unlock it and drop it back off all with the app.” Once finished, they were asked to record themselves repeating the same passage but this time in their “charismatic” voice. They were also asked to rate their own charismatic speech.

A total of 97 MTurk workers participated in our crowd-sourcing tasks. 60 raters’ birth gender was female, 36 raters male, and 1 preferred not to say. 68 of the raters were heterosexual, 11 were bisexual, and 16 were homosexual. 42 raters were attracted to females and 65 were attracted to males. The breakdown of the highest education level received by all raters was as follows: some school (1), high school (21), associates (19), BA (45), MA (10), PhD (1). The scores on the TIPI Big-Five personality dimensions range from 1 to 7 with a median of 4, while our raters’ average score was 5.12 for openness, 5.54 for conscientiousness, 3.70 for extroversion, 5.39 for agreeableness, and 4.91 for emotional stability. The raters’ personality distribution was skewed towards a higher score for the four personality dimensions except for extroversion.

Table 1: Pearson’s correlation for charisma and speaker traits.

Correlation	Speaker Traits
0.6 to 0.8	Liveliness, Enthusiasm, Persuasiveness Confidence
0.4 to 0.6	Extroverted, Eloquence, Trustworthiness Intelligence, Reasonableness
0.2 to 0.4	Sincerity, Fluency
-0.2 to -0.4	Coldness
-0.4 to -0.6	Boringness, Introversion, Weakness, Uncertainty, Ordinariness

4. Analysis and results

Using the voice clips, the ratings, and the raters’ information that we collected, we asked the following questions: How do raters define charisma in terms of other speaker traits? Does the genre of the recording influence charisma ratings? Does speaker gender influence raters’ charisma ratings or ratings on other speaker traits? What are the acoustic-prosodic and lexical properties of speech rated as charismatic? Do raters’ demographic information and personality characteristics influence their ratings? Does raters’ own speech correlate with their charisma ratings or their demographics/personality?

We used Pearson’s correlation, Krippendorff’s alpha, and paired t-tests to analyze the ratings of speaker traits, to identify the acoustic-prosodic and lexical characteristics of the rated voice clips, and to examine raters’ demographic and personality biases and assess their own speech data. We report significant results with a $p < 0.05$, unless otherwise stated.

4.1. Ratings of charisma and other speaker traits

Our 60 voice clips had an average charisma rating of 3.20 in range 1 to 5, indicating a fairly balanced dataset for charisma in speech. The least charismatic voice clip had an average rating of 1.53, and the most charismatic voice clip was rated at 4.50. To better understand raters’ definition of charisma using other potentially related speaker traits, we calculated Pearson’s correlations between ratings of charisma and ratings of the other speaker traits. The results are shown in Table 1, binned by 0.2 as suggested in Landis and Koch [21]. We also calculated the correlation for these traits separately for male and female speakers as well as male and female raters to see if there were differences in how charisma was defined gender-specifically but we did not find a statistically significant difference. Therefore, the definition of charisma in relation to the speaker’s other traits is consistent across both speaker genders and both rater genders.

For the inter-rater agreement, we calculated Krippendorff’s alpha over all speaker traits and obtained an alpha of 0.438, indicating reasonably good agreement among raters. Charisma was the sixth most agreed upon trait by our raters, with an alpha of 0.296. Our raters’ agreement on charisma ratings is comparable with previous work [12, 3, 13], which report alphas ranging from 0.22 to 0.31, depending on the quality of voice clips and the diversity of raters. The five most and the five least agreed-upon traits are shown in Table 2. It seems that higher activation traits are more agreed upon, and lower activation traits are more open to interpretation, which agrees with previous work [3].

4.2. Genre and charismatic ratings

Among the 60 clips we collected, 14 are interviews, 19 are educational lectures, and the other 27 are talks to more general audiences. We calculated the Pearson’s correlation for the charisma

Table 2: *Inter-rater agreement of speaker traits.*

Speaker Trait	α	Speaker Trait	α
Liveliness	0.389	Coldness	0.066
Enthusiasm	0.374	Reasonableness	0.132
Confidence	0.347	Ordinariness	0.133
Extroversion	0.297	Trustworthiness	0.153
Introversion	0.297	Fluency	0.157

ratings for each pair of genres, and we found that interviews are less charismatic than both educational lectures ($p = 0.009$) and talks ($p < 0.001$). However, talks and educational lectures are not rated significantly different on charisma. In these genres, when the speaker may be trying to make a point, they may seem more charismatic. For interviews, the goal of the genre may be more for factual transfer, so the speaker may appear less charismatic. This is consistent with previous works [3, 22, 14], in which speech genre and audience type were found to be significantly correlated with the charisma ratings.

4.3. Speaker gender and speaker trait ratings

We also examined whether speakers of different genders were rated as significantly different in charisma. While female speakers achieved a higher average charisma score than male speakers, the difference is not significant ($p = 0.153$). Male speakers were rated as less sincere ($p = 0.014$), less fluent ($p = 0.022$), and less extroverted ($p = 0.038$) than females, but more boring ($p = 0.001$) and more introverted ($p = 0.014$) using Pearson's correlation. A possible explanation is that 18 out of 27 talks were from females and talks were generally rated as more charismatic than other genres. The lower charisma score of males may be due to genre and not gender.

4.4. Acoustic-prosodic correlates of charisma

To study the acoustic properties of charismatic speech, we extracted 12 acoustic-prosodic features from each speaker clip, including the maximum, minimum, mean, and standard deviation of pitch and intensity, harmonics-to-noise ratio (HNR), jitter, shimmer, and speaking rate measured by the number of syllables per second. Although these features were extracted, we do not report maximum and minimum pitch and intensity because they provide similar interpretation as the standard deviation, in addition to being more susceptible to noise.

We examined the correlation over the acoustic-prosodic features and charisma scores to identify features that significantly indicate charisma. To account for the inherent difference in pitch between males and females, we normalized the mean pitch of males by 119 Hz with standard deviation 19 Hz and females by 210 Hz with standard deviation 27 Hz using mean values for American English speakers reported in Pépiot [23]. We found that mean intensity ($p = 0.013$), mean pitch ($p = 0.002$), speaking rate ($p = 0.001$), and variance in pitch ($p < 0.001$) were all positively correlated with charisma, meaning that voices that are louder, higher, faster, and with greater fluctuation in pitch were rated as more charismatic. We then considered whether there were any acoustic characteristics of charisma that were specific to speakers' gender. Once again, we calculated correlations of acoustic-prosodic features with charisma for each gender, without normalization. We observed a positive correlation with mean intensity ($p = 0.041$) and standard deviation in pitch ($p = 0.028$) for female speakers, and positive correlations with mean pitch ($p = 0.005$), speaking rate ($p = 0.011$) and standard deviation in pitch ($p = 0.001$) for male

speakers. So, not all acoustic-prosodic features of charisma that were found for all speakers were present within different genders. The mean intensity was only correlated with females' charismatic speech, while mean pitch and speaking rate were only correlated with males' charismatic speech. The correlation values of the acoustic-prosodic features ranged from 0.32 to 0.57, with mean intensity for all speakers having the lowest correlation value (0.32), and standard deviation in pitch for male speakers having the highest correlation value (0.57).

4.5. Lexical correlates of charisma

We extracted lexical features from the transcripts using Linguistic Inquiry and Word Count (LIWC) [24], for 73 categories such as affect words, social words, time orientation words, and words for cognitive, perceptual, and biological process. We calculated the correlation of charisma scores with these to see whether the perception of charisma is affected by the speech content.

The LIWC category of interrogative words ($p = 0.037$) was positively correlated with charisma, while first-person pronouns ($p = 0.017$), negative emotion words ($p = 0.014$), sadness words ($p = 0.002$), discrepancies ($p = 0.013$), and words of feeling ($p = 0.024$) were negatively correlated. This shows that speakers asking questions had high charisma ratings, while speakers who often referred to themselves and talked about their feelings, especially with negative emotion, received low ratings. We also examined gender-specific lexical correlates of charisma. For male speakers, religion words such as "faith" ($p = 0.041$) was positively correlated with charisma, while affect words ($p = 0.007$), positive emotion words ($p = 0.039$), negative emotion words ($p = 0.038$), sadness words ($p = 0.025$), and prepositions ($p = 0.028$) were negatively correlated. For female speakers, interrogative words ($p = 0.045$), numbers ($p = 0.048$), and words of seeing ($p = 0.026$) were positively correlated with charisma, while first-person pronouns ($p = 0.030$), words of feeling ($p = 0.018$), negative emotion words ($p = 0.036$), sadness words ($p = 0.006$), words describing cognitive processes ($p = 0.047$), and discrepancies ($p = 0.022$) were negatively correlated. By comparing the lexical correlates of male and female charisma, we see that there are some differences but also some shared characteristics: speakers that use negative emotional words were rated as less charismatic regardless of gender. The absolute correlation values of lexical features ranged from 0.27 to 0.40 when considering all speakers; for gender-specific groups the absolute values were generally higher, ranging from 0.36 to 0.49. This indicates that we may have a better interpretation of charisma when taking gender into account.

4.6. Raters' characteristics and their speaker ratings

Focusing next on the rater's side, we examined their demographics and personalities to see whether a rater's birth gender, gender attraction, education level, and personality scores influence how they rate speaker traits.

To determine whether a rater's birth gender influenced their ratings of speaker traits, we calculated Pearson's correlation of raters' gender and trait ratings. We found that male raters rated speakers in general as weaker ($p = 0.015$) and colder ($p = 0.001$) than female raters did. We also examined whether the birth gender influenced how they rated speakers of different genders by calculating the correlation of rater's gender and ratings on males and on females separately. When judging male speakers, male raters rated them as weaker ($p = 0.019$) and less fluent ($p = 0.040$) than female raters did. For female speakers, male raters rated them as colder ($p = 0.003$), more introverted ($p =$

0.022) and less extroverted ($p = 0.015$) than female raters did. In addition to birth gender, we were also interested in seeing whether raters rated speakers whose gender they are *attracted to* differently. We found that raters judged the attracted gender as more introverted ($p < 0.001$) and boring ($p = 0.032$), and less confident ($p = 0.042$), extroverted ($p = 0.006$), trustworthy ($p = 0.046$), reasonable ($p = 0.037$), and charismatic ($p = 0.020$). This might be because a majority of our raters happened to be heterosexual female and the voice clips with male speakers were generally rated as less charismatic, as noted above 4.3. We next studied the correlation between raters' education level and speaker trait ratings. We found that the higher their education level, the less ordinary ($p < 0.001$), boring ($p = 0.017$), intelligent ($p = 0.039$), and fluent ($p = 0.015$), and the more eloquent ($p = 0.007$) and lively ($p = 0.044$) they rated speakers. This suggests that raters may use themselves as a reference when judging the speakers.

To compare raters' personalities to their ratings of speaker traits, we calculated Pearson's correlation between the raters' TIPI personality scores and their trait ratings. Raters with higher scores in openness, conscientiousness, agreeableness, and emotional stability, tended to rate speakers higher in charisma and in traits that positively correlated with charisma, but lower in traits that negatively correlated with charisma. This suggests that raters may project some of their own personalities in rating others. However, raters with higher personality scores in extroversion tended to rate speakers lower in charisma and in traits positively correlated with charisma, while higher in traits negatively correlated with charisma. This could be explained by the correlations between personality and self charisma rating, described below in 4.7, in which extroversion was positively correlated with self charisma scores. The more extroverted the raters are, the higher they assessed themselves in charisma, and thus perhaps the lower they evaluated other speakers' charisma.

The absolute correlation values of the raters' characteristics and their ratings were fairly weak, ranging from 0.06 to 0.20. Although the rater's own characteristics had some influence on their ratings, the effect was weaker than the characteristics of the speaker's speech.

4.7. Analysis of rater's speech

We also examined how raters adjusted their speech when asked to speak charismatically and how raters' own speech may have influenced how they rated other speakers. We analyzed differences in speaking style when raters were asked to speak normally or when asked to speak the same text charismatically and compared these to their demographics and personality.

We calculated raters' speaking differences or *adjustment* as the change in acoustic features from each rater's natural speech to their charismatic speech, measured by paired t-tests. Compared with their natural speech, raters increased their mean intensity ($p < 0.001$), mean pitch ($p < 0.001$) and standard deviation of pitch ($p < 0.001$), and decreased their HNR ($p = 0.028$) when asked to be charismatic. This suggests that the raters' own adjustments were similar to how they rated the speakers' voice clips, except that they lowered their own HNR in charismatic speech but did not apparently judge the speakers' charisma by HNR. For gender-specific rater groups, we found that female raters increased their mean intensity ($p < 0.001$), similar to the acoustic correlates of charisma for the female voice clips. Male raters increased their mean pitch ($p < 0.001$) and speaking rate ($p = 0.025$) as we found in the male voice clips, but they also increased mean intensity ($p < 0.001$), jitter ($p = 0.025$) and de-

creased their HNR ($p = 0.015$). The overall trend shows that raters do change their voices based on what they believe sounds more charismatic, but they also increase other acoustic features they may be less aware of when rating others.

When we compared the charisma adjustment between male and female raters by calculating the Pearson's correlation between the raters' adjustment and the raters' birth gender, we found that males had a higher positive difference in mean pitch ($p < 0.001$), speaking rate ($p = 0.012$), and variance of their pitch ($p < 0.001$) than females. Furthermore, if we look at rater adjustment compared with how raters judged their own charismatic voices, we find no significant results for females or for a combination of both genders; however, males increase their variation in pitch ($p = 0.049$) the more charismatic they think they are. This suggests that males exaggerate the features we found to be associated with charisma more than females do when producing charismatic speech, and that male raters who see themselves as more charismatic tend to exaggerate their charismatic speech even more.

The education level of a rater had no effect on their charisma adjustment, while personality had a slight impact. The higher a rater's extroversion score was, the more they increased the variance in their pitch ($p = 0.037$). Moreover, raters with higher agreeableness had a lower increase in their mean pitch ($p = 0.001$), and raters with higher emotional stability had a slightly higher positive difference in their speaking rate ($p = 0.012$). This suggests raters with higher agreeableness may be less charismatic since they decrease the acoustic features associated with charisma, while those with higher extroversion and higher stability may be more charismatic. We also examined whether rater's personality impacted how they rated their own voice. We found that raters scored themselves higher on charisma if they had a higher openness ($p = 0.008$), conscientiousness ($p < 0.001$), extroversion ($p < 0.001$), or agreeableness ($p = 0.037$) scores. This trend is also true for both conscientiousness and extroversion scores for females ($p = 0.001$, 0.011) and males ($p = 0.033$, 0.003) when we separate by gender. It is interesting to note that, although both openness and conscientiousness had no impact on raters' adjustment to producing charismatic speech acoustically, they did have an impact on their charisma self-ratings.

5. Conclusions and future work

In this research on the role of gender, demographics, and personality in the production and perception of charisma, we studied the definition of charisma by identifying other speaker traits that positively or negative correlated with charisma. We analyzed acoustic-prosodic correlates of charisma and found that charismatic voices were louder, higher, faster, with greater variation in pitch, although there was some difference between male and female charismatic voices. Text-based correlates of charisma showed that speakers who used more questions were rated as more charismatic, while speakers who talked about themselves and their feelings, especially conveying negative emotions were rated as less charismatic regardless of their gender. We also found differences in the way raters with different demographics and personalities rated speakers for charisma and other speaker traits. These findings reveal significant individual differences that should be identified and taken into account in future research. Also, while prior work has shown that culture affects charisma ratings on politician speech [4, 11], we do not know whether this also applies to non-political charismatic speech ratings; we will explore this in future research.

6. References

- [1] M. Weber, *The theory of social and economic organization*. Simon and Schuster, 2009, based on *The theory of social and economic organization*, Max Weber; A M Henderson; Talcott Parsons New York, Oxford University Press, 1947, a translation of part I of Max Weber's *Wirtschaft und Gesellschaft*.
- [2] A. Rosenberg and J. Hirschberg, "Acoustic/prosodic and lexical correlates of charismatic speech," in *Interspeech 2005*, Lisbon, 2005.
- [3] —, "Charisma perception from text and speech," *Speech Communication*, vol. 51, no. 7, pp. 640–655, 2009.
- [4] F. Biadys, A. Rosenberg, R. Carlson, J. Hirschberg, and E. Strangert, "A cross-cultural comparison of american, palestinian, and swedish perception of charismatic speech," in *4th International Conference on Speech Prosody 2008*, Campinas, Brazil, 2008, pp. 579–582.
- [5] R. Signorello, F. D'Errico, I. Poggi, D. Demolin, and P. Mairano, "Charisma perception in political speech: a case study," in *International Conference on Speech and Corpora (GSCP 2012)*, 2012, pp. 343–348.
- [6] R. Signorello, F. Derrico, I. Poggi, and D. Demolin, "How charisma is perceived from speech: A multidimensional approach," in *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing*, 2012, pp. 435–440.
- [7] S. Scherer, G. Layher, J. Kane, H. Neumann, and N. Campbell, "An audiovisual political speech analysis incorporating eye-tracking and perception data," in *Language Resources and Evaluation Conference 2021 (LREC)*, 2012, pp. 1114–1120.
- [8] L. Chen, G. Feng, J. Joe, C. W. Leong, C. Kitchen, and C. M. Lee, "Towards automated assessment of public speaking skills using multimodal cues," in *Proceedings of the 16th International Conference on Multimodal Interaction*, 2014, pp. 200–203.
- [9] O. Niebuhr and J. Michalsky, "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 *International Conference on Computational Science and Its Applications*, 2019, pp. 375–390.
- [10] O. Niebuhr, A. Brem, and S. Tegtmeier, "Advancing research and practice in entrepreneurship through speech analysis – from descriptive rhetorical terms to phonetically informed acoustic charisma profiles," *Journal of Speech Sciences*, vol. 6, no. 1, pp. 3–26, 2017.
- [11] F. D'Errico, R. Signorello, D. Demolin, and I. Poggi, "The perception of charisma from voice: A cross-cultural study," in *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, 2013, pp. 552–557.
- [12] A. Cullen, A. Hines, and N. Harte, "Building a database of political speech: Does culture matter in charisma annotations?" in *Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge*, 2014, pp. 27–31.
- [13] F. Weninger, J. Krajewski, A. Batliner, and B. Schuller, "The voice of leadership: Models and performances of automatic analysis in online speeches," *IEEE Transactions on Affective Computing*, vol. 3, no. 4, pp. 496–508, 2012.
- [14] H. Mixdorff, O. Niebuhr, and A. Hönemann, "Model-based prosodic analysis of charismatic speech," in *Proceedings of the 9th International Conference on Speech Prosody*, Poznan, Poland, 2018, pp. 1–5.
- [15] O. Niebuhr and S. Gonzalez, "Do sound segments contribute to sounding charismatic? evidence from a case study of steve jobs' and mark zuckerberg's vowel spaces," *International Journal of Acoustics & Vibration*, vol. 24, no. 2, pp. 343–355, 2019.
- [16] S. Berger, O. Niebuhr, and B. Peters, "Winning over an audience—a perception-based analysis of prosodic features of charismatic speech," in *Proceedings of the 43rd Annual Conference of the German Acoustical Society*, Kiel, Germany, 2017, pp. 1454–1457.
- [17] E. Novák-Tót, O. Niebuhr, and A. Chen, "A gender bias in the acoustic-melodic features of charismatic speech?" in *Interspeech 2017*, Stockholm, 2017, pp. 2248–2252.
- [18] O. Niebuhr, S. Tegtmeier, and T. Schweisfurth, "Female speakers benefit more than male speakers from prosodic charisma training — a before-after analysis of 12-weeks and 4-h courses," *Frontiers in Communications*, April 2019.
- [19] S. D. Gosling, P. J. Rentfrow, and W. B. Swann, "A very brief measure of the big-five personality domains," *Journal of Research in Personality*, vol. 37, no. 6, pp. 504–528, 2003.
- [20] R. R. McCrae and P. T. Costa, "Validation of the five-factor model of personality across instruments and observers," *Journal of Personality and Social Psychology*, vol. 52, no. 1, pp. 81–90, 1987.
- [21] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, pp. 159–174, March 1977.
- [22] O. Niebuhr, J. Voße, and A. Brem, "What makes a charismatic speaker? a computer-based acoustic-prosodic analysis of steve jobs tone of voice," *Computers in Human Behavior*, vol. 64, pp. 366–382, 2016.
- [23] E. Pépiot, "Male and female speech: a study of mean f0, f0 range, phonation type and speech rate in parisian french and american english speakers," in *Speech Prosody 7*, 2014, pp. 305–309.
- [24] J. W. Pennebaker, R. L. Boyd, K. Jordan, and K. Blackburn, "The development and psychometric properties of liwc2015," University of Texas at Austin, Tech. Rep., 2015.