



The acoustic fingerprint of a charismatic voice - Initial evidence from correlations between long-term spectral features and listener ratings

Oliver Niebuhr¹, Radek Skarnitzl², Lea Tylečková²

¹SDU Electrical Engineering, Mads Clausen Institute, University of Southern Denmark, Sønderborg

²Institute of Phonetics, Faculty of Arts, Charles University, Czech Republic

oniebuhr@mci.sdu.dk, radek.skarnitzl@ff.cuni.cz, lea.tyleckova@seznam.cz

Abstract

What makes a charismatic speaker? The present study extends this question into the prosodic dimension of voice quality. We analyzed various F0, LTAS and LTF long-term spectral characteristics from 12 L2 speakers of English who were recorded while giving entrepreneurial speeches. The results of the acoustic analysis were correlated with indirect judgments of the entrepreneurs' charismatic performances by 98 listeners. The correlations we found replicate previous findings in that a larger F0 range and a higher/lower F0 level are beneficial for a male/female speaker's perceived charisma. Moreover, LTAS settings that are indicative of a fuller and less breathy voice also led to higher speaker charisma ratings. The same applies to LTF settings that are indicative of a larger body or vocal-tract size. The findings are discussed with respect to their implications for measuring and training charismatic speech, traditional rhetoric statements and the definition of charisma.

Index Terms: Charisma, LTF, LTAS, F0, entrepreneur, voice.

1. Introduction

Understanding the phenomenon of charisma is of "immense importance for society [...] because charismatic leaders wield enormous power and can use this for great good or evil" [1: 294]. Charisma is not a mysterious gift reserved for a few chosen people [2] but a learnable, improvable skill [3,4]. Charismatic speech is relevant to many everyday situations. Being charismatic can result in more fruitful speed-dating and brainstorming outputs [5], makes addressees (e.g., students) learn better or more [6,3], increases the chance of attracting investors or raising start-up funding [7], makes a product or service more credible and likable to customers [8], and functions as a major career catalyst [9,10].

Considering what is known about *who* can be charismatic and *what* charismatic speakers can achieve with their skills, still relatively little is known about *how* charismatic speech manifests itself in the speech signal and *how* signal parameters have to change in order to make a speaker sound more charismatic in the ears of listeners. Researchers with backgrounds in rhetoric, management, and (social) psychology have shed some light on these questions in the recent past, e.g., [11-16]. However, the descriptive rhetorical labels they work with – such as "rich", "animated", "fluent", and "durable" – are hard to operationalize and replicate experimentally, and their instructive value for trainers and learners of charismatic speech is strongly limited, see [17] for an example.

Compared to the metaphorical or descriptive charisma labels of rhetoric, the digital speech-signal processing and analysis techniques of modern phonetics allow a much more fine-

grained approach and give us phoneticians the particular opportunity to turn the supposed elusive mystery of charismatic speech into an objective and tangible research subject. Rosenberg and Hirschberg [18,19] were the first who called for an empirical definition of speaker charisma and provided a first answer to their call by projecting ratings of perceived charisma attributes onto acoustic-prosodic features of speakers. This empirical foundation was subsequently further supported and enriched by many similar studies [20-30].

When people think of charisma, they virtually always think of politicians. Most popular-science articles or guidebooks on charisma begin with a portrait of, e.g., Martin Luther King, Jr., Winston Churchill, Margaret Thatcher, Jacques Chirac, Bill Clinton, or Barack Obama; and, in fact, many previous scientific papers also revolved around these or similarly prominent political figures. This is probably because the original definition of charisma by Weber [2] saw a societal crisis as a prerequisite for this communicative skill to come to light; and leading a whole society through difficult times is obviously the domain of politics.

Our focus, however, lies on economic key players. This is firstly due to the fact that, nowadays, the boundaries between politics and economics become blurred. With growing financial resources, companies gain a significant influence on politics and their CEOs are no longer just quiet puppet masters in the background. Rather, more and more modern CEOs become visible figureheads of their companies and, thus, part of their companies' marketing strategies and brand images. Popular and particularly outgoing examples of this "species" of CEOs are Steve Jobs, Mark Zuckerberg, Dieter Zetsche, Elon Musk, and Oprah Winfrey. Secondly, our focus on business rather than political players takes into account the fact that modern companies and whole national economies are faced with a global competition of ideas and innovations. Entrepreneurs play a large and still growing role in this context. In fact, they have become a pillar of economic growth [31], and High-Tec societies are promoting such start-up initiatives in a significant way and from different angles. Being unknown when founding a new business, entrepreneurs need to give priority to legitimizing their activities [32,33]. One important form of these activities is the so-called "investor pitch", i.e. a short overview presentation of one's business idea or plan for a group of potential investors and/or decision-makers, sometimes organized as large "pitching contests". In order to be successful in such a framework, it is crucial to be persuasive – and the major vehicle of persuasiveness, perhaps even more important than the idea/plan itself [6,11,12], is the entrepreneur's prosody.

Against this background, we have set up a trans-disciplinary line of research called PERCY (Persuasiveness

and Creativity¹). One of PERCY's initial key findings was that the prosodic profiles which make politicians sound more charismatic [18-24,27,30] also apply to CEOs in general and entrepreneurs in particular. Moreover, we increased the tangibility and accessibility of charismatic speech by adding to the known prosodic charisma profile (which was until then mainly based on F0, intensity, and speaking rate [18-24]) parameters of rhythm, emphatic accentuation, pitch-accent shape and timing, as well as patterns of pausing, disfluency, and speech-reduction.

In the present paper, we further extend our understanding of the prosodic charisma profile in the direction of *voice quality*, i.e. long-term spectral characteristics of speakers. Unlike F0, which we include here as a basic, well-studied feature, voice quality has so far barely been addressed in acoustic-phonetic charisma research, as it is hard to measure and still in the process of being standardized (e.g., [34]) and because it requires a constant and high recording quality, which is difficult to obtain for ecologically valid speeches of, e.g., popular speakers, who were recorded for various purposes and under various conditions.

Therefore, our data material was specifically recorded for the purpose of phonetic analyses. It consists of investor-pitch speeches that were given by real entrepreneurs to a peer-group audience under acoustically controlled conditions. Like in previous studies of other researchers [18,19,21] (and inter alia), we correlated the voice-quality measurements with naive listener ratings on the speakers' performances in order to draw conclusions on which voice-quality measures are involved in perceived speaker charisma and how. The conclusions drawn here are of a preliminary nature, as our set of entrepreneur speeches is still limited but constantly growing. Thus, we see the primary aim of the present study in identifying those parameters, results, and rating questions that are worth investigating in greater detail with a larger data set in future studies.

2. Method

The speech corpus we analyzed included investor pitches given by 12 post-graduate students (8 males, 4 females) of the Dept. of Technology Entrepreneurship and Innovation at the University of Southern Denmark (www.sdu.dk/tei). The young entrepreneurs were recorded while presenting (i.e. "pitching") their business ideas to a group of business-engineering students. All 12 young entrepreneurs were L2 speakers of English. Most of them (7) had a Western Germanic L1 background (German or Danish), two speakers had Mandarin Chinese as their native language, another two were native speakers of Urdu, and one had an L1 background in French. The speakers' average age was 23.1 years.

The recordings took place in the sound-treated SDU MCI Innovation Lab at the University of Southern Denmark in Sønderborg², see Figure 1. The speakers were standing during the recording, yet their position relative to the microphone remained roughly constant. The speech was recorded digitally at a 44.1 kHz/24-bit quality with a Zoom H4 device. It was connected to an array of pressure-zone microphones (Sennheiser MEB 114) that are subtly embedded in the floor of the Innovation Lab.



Figure 1: *The investor-pitch recording situation within the SDU MCI Innovation Lab in Sønderborg.*

We used 40–50 second excerpts (mean duration = 46.4 s) for our analyses, cut out at major prosodic-phrase boundaries from the middle of each investor pitch. The acoustic analysis itself included three signal domains: fundamental frequency (F0), long-term average spectrum (LTAS), and long-term formant distribution (LTF). As our study was exploratory in nature, we measured multiple potentially relevant/redundant parameters per signal domain.

For F0, central tendency and variability indicators were extracted in a 60–450 Hz range using autocorrelation in Praat (v.6.036, www.praat.org). The central tendency was represented by standard measures like mean and median. Additionally, the F0 baseline was calculated as it is supposed to represent a speaker's neutral F0 level better than mean and median and to be more robust against different technical and behavioral conditions [35,36]. The baseline corresponds to the 7.64th percentile of F0 values. Variability indicators included standard deviation, coefficient of variation (varco) and the percentile range (i.e., the difference between the 10th and 90th percentile).

The LTAS analysis was conducted with the pitch-corrected algorithm [37] implemented in Praat. The sound files were resampled to 16 kHz to that end. Three traditional LTAS parameters were measured: (i) the alpha measure [38], which corresponds to the ratio of energy between the 0–1 kHz and 1–5 kHz and reflects spectral slope; (ii) the L1–L0 value, which corresponds to the energy difference between the F1 (300–800 Hz) and F0 (50–300 Hz) regions (see, e.g., [39,40]) and pertains to the mode of phonation; and (iii) the energy ratio between 1–5 and 5–8 kHz (e.g., [40]), which reflects the degree of breathiness of the speech signal.

Voice quality has also been associated with the so-called speaker's formant (SF), i.e. a strong acoustic energy peak. The peak is located between 3 and 4 kHz in the LTAS of professional male voices. The stronger it gets the more it makes a (male) voice sound full and sonorous. In accordance with [41], the strength of the SF was assessed by calculating the energy difference between frequency range of 300-800 Hz (L1) on the one hand and the ranges of 2–3 kHz and 3–4 kHz on the other; these additional two LTAS-based SF measures are referred to here as L1-(2–3) and L1-(3–4).

For LTF [42], we first created chains of vowels and non-nasal sonorant consonant sounds (see [43]) and then extracted

¹ https://www.sdu.dk/da/om_sdu/institutter_centre/sdu+electrical+engineering/researchprojects/percy

² <https://www.youtube.com/watch?v=yB8NRyGfM0Q>

F1–F3 values every 10 ms using sex-specific default settings in Praat. Subsequently, we calculated spectral moments (the mean, standard deviation, skewness, and kurtosis) of the three distributions for each speaker.

The measured parameter values of our 12 speakers were, for each parameter, correlated with average listener ratings of the speakers’ pitching performances. To that end, an online survey was created that presented the 12 investor-pitch excerpts in an individually randomized order and in the context of two questions: (i) How likely is it that you would dare to invest some money into the speaker’s company? (ii) How much management/leadership experience do you think the speaker has? A total of 98 listeners completed the online survey. They rated the investor-pitch excerpts in percentages (0-100) for question (i) and in years (0-10 or more) for question (ii). The average ratings that we received for the two questions on each speaker were correlated with the speaker’s individual acoustic measurements using Pearson’s product moment correlation coefficient.

3. Results

The number of correlated values corresponds to the number of speakers, i.e. $n=12$. Due to the relatively small number of values, and in view of the fact that the present study is primarily meant to lay the foundation for a follow-up analysis with a larger dataset, we decided to lower the significance threshold from $p \leq 0.05$ to $p \leq 0.1$. The lowered threshold avoids that we prematurely exclude parameters that are weakly but nevertheless systematically correlated with speaker-performance ratings and whose statistical significance will therefore come out more clearly in combination with a larger dataset. For $n=12$ ($df=10$), the threshold of $p \leq 0.1$ is reached at correlation coefficients larger than 0.5 ($r \geq 0.496$).

Based on this threshold of $r \geq 0.5$, we found a number of statistically significant correlations between acoustic measurements and listener ratings. Most of these correlations occurred in combination with the speaker-performance ratings based on question (i), which asked for the likelihood of the listeners to invest money in the speaker’s company. That most correlations emerged for this question is probably due to the fact that it was easier to answer than question (ii). Unlike the latter, question (i) asked listeners to make judgments about themselves and not about the speaker they listened to.

3.1 Fundamental frequency parameters

Correlations between speaker-performance ratings (question (i)) and two F0 dispersion parameters came out clearly significant. The 98 listeners were more likely to invest in the company of a speaker if his/her speech was characterized by a more variable speech melody in terms of a higher F0 standard deviation ($r=0.64$, $p=0.031$) and higher F0 percentile range ($r=0.69$, $p=0.013$).

In addition, we found significant correlations between question (i) speaker performance ratings and all three measurements of F0 level, i.e. mean F0 ($r=-0.72$, $p=0.008$), median F0 ($r=-0.73$, $p=0.006$), and baseline F0 ($r=-0.63$, $p=0.037$). However, these correlations were negative, i.e. the speakers’ performances increased with decreasing F0 levels. As this result is inconsistent with conclusions drawn from previous studies (see [26] for a summary), we examined the corresponding correlations further and, indeed, found them to be an artefact of speaker sex. That is, the correlations actually reflect two other facts: First, listeners were overall less likely

to invest in female speakers’ companies (42.3% likelihood of listener investment as compared to 54.8% for the male speakers’ companies); and second, male speakers performed better in the ears of listeners when they had a higher F0 level ($\approx +60$ Hz in the F0 baseline caused a 7.1% increase in listeners’ investment likelihood), whereas female speakers performed better with a lower F0 level (≈ -12 Hz in the F0 base-line led to a 15.4% increase in investment likelihood). Note that there were no similar speaker-sex differences for the F0 dispersion measures and their correlations with listener ratings.

3.2 Traditional LTAS parameters

Three significant correlations were found for the traditional LTAS parameters. They all point in the same direction, i.e. the louder and less breathy a speaker’s voice was, the better the speaker’s performance was rated by the listeners. This fact applies independently of speaker sex and manifests itself in the form of a positive correlation between the alpha ratio and the speaker’s estimated years of management/leadership experience ($r=0.51$, $p=0.085$), as well as in a negative correlation between the 1-5/5-8 kHz spectral-energy ratio and both question ratings, i.e. investment likelihood ($r=-0.50$, $p=0.091$) and the speaker’s estimated years of management/leadership experience ($r=-0.68$, $p=0.016$), see Figure 2. The two negative correlations are notably steep. In consequence, a decrease in the 1-5/5-8 spectral-energy ratio by about 25% increases the likelihood to invest in the speaker’s company on average by 70-100% and likewise almost doubles the speaker’s estimated management/leadership experience. A change in the 1-5/5-8 spectral-energy ratio of about 25% is actually not that much (and thus achievable through training) insofar as both different speakers (of the same sex) and different emotional states can cause 1-5/5-8 spectral-energy differences that are much larger than 25%, see, e.g., [40]. The L1-L0 measure was not correlated with any speaker-performance ratings.

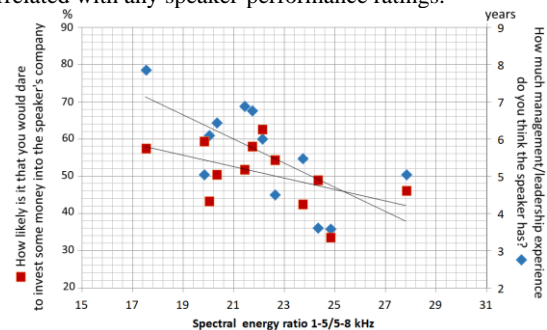


Figure 2: Correlations between 1-5/5-8 kHz ratios and speaker-performance ratings based on Q(i) and Q(ii).

3.3 Speaker’s formant (SF) parameters

No clear pattern emerged for female voices (even when the frequency ranges given in section 2 were multiplied by 1.2 to take into account the, on average, shorter female vocal tract). However, the parameters reflecting the SF in male speakers yielded the strongest correlations found in this study. As is shown in Figure 3, this concerns the correlation of the listeners’ willingness to invest in the speakers’ business idea with both L1-(2–3) ($r=-0.74$, $p=0.037$) and L1-(3–4) ($r=-0.81$, $p=0.015$). The two negative correlations mean that listeners’ investment likelihood rose when the SF (reflected in the two frequency ranges) increased in amplitude and, thus, reduced the difference to the minuend L1.

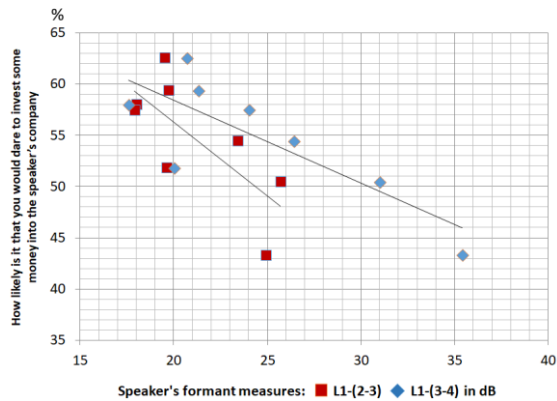


Figure 3: Correlations between the two SF measures and speaker-performance ratings based on $Q(i)$.

3.4 LTF parameters

The LTF analysis yielded a significant correlation between the standard deviation of the third formant (F3) distribution and the listeners' willingness to invest their money into a speaker's company ($r=-0.52$, $p=0.077$).

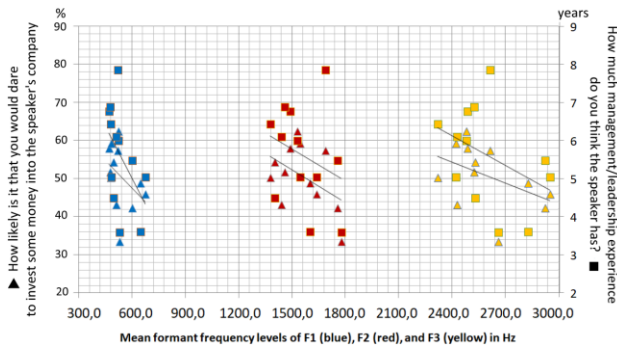


Figure 4: Correlations between F1-F3 and speaker-performance ratings based on $Q(i)$ and $Q(ii)$.

There are no other significant correlations between speaker-performance ratings and formant parameters. However, despite not being significant, one group of correlation coefficients still stands out; they are *all* consistently negative, *all* concern the mean levels of the three formants, and *all* are nearly significant, with r lying between -0.47 and -0.39 . So, if these individual correlations are integrated – either across question (i)-(ii) or across the three formants of each question (see [44] for the mathematical details of this integration) – they actually represent a very significant finding whose alpha-error levels are below $p<0.01$. That is, what we found is basically that the lower a speaker's mean formant frequencies are the more likely is the listener to invest in his/her company and the higher s/he perceives the speaker's level of management/leadership experience to be, see Figure 4. This was true for the stimuli of both male and female speakers.

4. Discussion and conclusions

This paper asked the question which long-term spectral characteristics make a speaker sound (more) charismatic. Based on a pilot dataset of 12 speakers, we took several established LTAS and LTF measures and also included F0 so as to link the present study to previous research in which the relevance of F0 has been documented. The measurements

from the pilot dataset were correlated with perceived speaker charisma, represented here by two key abilities to (i) persuade listeners [19] and (ii) gather followers [1] that were indirectly judged by 98 naive listeners in terms of (i) the likelihood to invest own money into the speaker's company/idea and (ii) the estimated management/leadership experience of the speaker.

Our results on F0 are consistent with previous studies in that they show that more extensive and variable pitch patterns add to a speaker's perceived charisma. Moreover, we show that the "F0 baseline" (7.64th percentile) and the "F0 percentile range" (90th -10th percentile) represent insightful and robust alternatives to the more simple classic measures of the F0 mean/median and the F0 range. Our F0 results also support the conclusion drawn by [28] that female speakers have to lower rather than raise their F0 level and are, everything else being equal, perceived to be less charismatic than male speakers.

The replication of previous F0 findings supports the general validity of our results. In this light, we draw the following conclusions on our LTAS and LTF data. The energy ratio between 1-5 and 5-8 kHz [42] is a particularly fruitful voice-quality parameter for measuring perceived speaker charisma. In combination with the less fruitful alpha ratio, it suggests that louder, less breathy, and fuller voices boost speaker charisma. The results on the speaker's formant (SF) are in line with this conclusion. The L1-L0 measure can be omitted in future studies. For LTF, our results suggest a link between charisma and overall lower formant levels. We take this as evidence that it is a larger body or vocal-tract size that makes a speaker sound more charismatic (this could be one important reason for the general female charisma disadvantage). Note that all of these LTAS and LTF findings perfectly agree with what is stated, in descriptive terms, in rhetoric studies and guidebooks for a long time [45]. We just put these statements on a quantitative empirical basis. The relevance of the third formant's standard deviation is harder to make sense of. Our preliminary explanation is that less lip activity – perhaps in the form of a constant smile – adds to perceived charisma.

In summary, we can conclude that the acoustic fingerprint of a charismatic voice is multifarious and complex, but, unlike that of F0, probably independent of speaker sex. Voice-quality measures have to be(come) an integral part of acoustic-phonetic charisma analyses; and training a speaker's voice seems to be a very rewarding investment given how strongly even relatively small parameter changes already affected the listeners' perception of a speaker. We will proceed with repeating our analysis on a larger set of speakers and, in this context, also address the question whether our findings are directly or indirectly related to voice quality, in the latter case via changes in the speaker's perceived emotional state. Finally, in view of the obvious differences between our findings on perceived charisma and previous findings on perceived attractiveness [29], we will also explore how our measures can contribute to clearer separate definitions of the two concepts. The complexity of the charisma concept alone is reflected in the present study in the fact that the listeners' ratings on the two questions $Q(i)$ and $Q(ii)$ were not significantly correlated with each other. That is, studies determining perceived charisma should always include multiple attributes or scales, as is suggested in [19]. In fact, determining how to properly capture perceived charisma is a research topic in its own right.

Acknowledgement: The second author was supported from European Regional Development Fund-Project "Creativity and Adaptability as Conditions of the Success of Europe in an Interrelated World" (No. CZ.02.1.01/0.0/0.0/16_019/0000734).

5. References

- [1] Antonakis J., Bastardoz N., and Jacquart P., "Charisma: An ill-defined and ill-measured gift," *Annual Review of Organizational Psychology & Organizational Behavior* 3, pp. 293-319, 2016.
- [2] Weber, M., *The Theory of Social and Economic Organization*. New York: Free Press, 1947.
- [3] Towler A.J., "Effects of charismatic influence training on attitudes, behavior, and performance," *Personnel Psychology* 56, pp. 363-381, 2003.
- [4] Antonakis J., Fenley M., Liechti S., "Can charisma be taught? Tests of two interventions," *Acad. Manag. Learn. Educ.* 10, pp. 374–396, 2011.
- [5] Pentland A., *Honest Signals - How they shape our world*. Cambridge: MIT Press, 2008.
- [6] Shea C. and Howell J.M., "Charismatic leadership and task feedback: A laboratory study of their effects on self-efficacy and task performance," *The Leadership Quarterly* 10, pp. 375-396, 1999.
- [7] Davis B.C, Hmieleski K.M, Webb J.W, Coombs J.E., "Funders' positive affective reactions to entrepreneurs' crowd-funding pitches: The influence of perceived product creativity and entrepreneurial passion," *J. of Bus. Vent.* 32, pp. 90–106, 2017.
- [8] Gélinas-Chebat C., Chebat J.C., Vaninsky A., "Voice and Advertising : Effects of Intonation and Intensity of Voice on Source Credibility, Attitudes and the Intend to Buy," *Perceptual and Motor Skills* 83, pp. 243-262, 1996.
- [9] Bodow S., "Charmed I'm sure," *USA Today*, 2012.
- [10] Jacquart P., Antonakis J., "When does charisma matter for top-level leaders? Effect of attributional Ambiguity," *Acad. Manag. J.* 58, pp.1051–1074, 2015.
- [11] Holladay S.J., Coombs W.T., "Communicating visions: An exploration of the role of delivery in the creation of leader charisma," *Man. Com. Quart.* 6, pp. 405-427, 1993.
- [12] Holladay S.J., Coombs W.T., "Speaking of visions and visions being spoken an exploration of the effects of content and delivery on perceptions of leader charisma," *Man. Com. Quart.* 8, pp. 165-189, 1994.
- [13] Den Hartog D.N. and Verbarg R.M., "Charisma and Rhetoric: Communicative techniques of international business leaders," *The Leadership Quarterly* 8, 355-391, 1997.
- [14] Cyphert D., "The rhetorical analysis of business speech - Unresolved questions," *J. of Bus. Com.* 47, 346-368.
- [15] Antonakis J., Fenley M., Liechti S., "Learning charisma: Transform yourself into someone people want to follow," *Harvard Bus. Rev.*, pp. 127–130, 2012.
- [16] Sørensen, L.S., "How to Grow an Apple: Did Steve Jobs Speak Apple to Success?," MA thesis, Aalborg Univ., DK, 2013.
- [17] Niebuhr O., Tegmeier S., Brem A., "Advancing research and practice in entrepreneurship through speech analysis – from descriptive rhetorical terms to phonetically informed acoustic charisma metrics," *J. of Speech Sciences* 6, pp. 3–26, 2017.
- [18] Rosenberg A. and Hirschberg J., "Acoustic/Prosodic and lexical correlates of charismatic speech," *Proc. Eurospeech 2005*, Lisbon, Portugal, pp. 1–4, 2005.
- [19] Rosenberg A. and Hirschberg J., "Charisma perception from text and speech," *Speech Com.* 51, pp. 640–655, 2009.
- [20] Touati P., "Prosodic aspects of political rhetoric," *Proc. ESCA Workshop on Prosody*, Lund, Sweden, pp. 168–171, 1993.
- [21] Biadsy F., Rosenberg A., Carlson R., Hirschberg J., Strangert E., "A cross-cultural comparison of American, Palestinian, and Swedish perception of charismatic speech," *Proc. Speech Prosody*, Campinas, Brazil, pp. 579–582, 2008.
- [22] Signorello R., D'Errico F., Poggi I., and Demolin D., "How charisma is perceived from speech. A multidimensional approach," *Proc. ASE/IEEE international conference on social computing*, Amsterdam, Netherlands, pp. 435-440, 2012.
- [23] Signorello R., D'Errico F., Poggi I., Demolin D., and Mairano P., "Charisma perception in political speech: A case study," *Proc. International conference on speech and corpora*, Belo Horizonte, Brazil, 343-348, 2012.
- [24] D'Errico F., Signorello R., Demolin D., Poggi I., "The perception of charisma from voice. A crosscultural Study," *Proc. Humaine Association Conference on Affective Computing and Intelligent Interaction*, Geneva, Switzerland, pp. 552–557, 2013.
- [25] Niebuhr O., Brem A., Novák-Tót E., "Prosodic constructions of charisma in business speeches – A contrastive acoustic analysis of Steve Jobs and Mark Zuckerberg," in: *Proc. 8th International Conference of Speech Prosody*, Boston, USA, 2016, pp. 1–3.
- [26] Niebuhr O., Voße J., and Brem, A., "What makes a charismatic speaker? A computer-based acoustic prosodic analysis of Steve Jobs tone of voice," *Computers and Human Behavior* 64, pp. 366–382, 2016.
- [27] Hiroyuki T., Rathcke T., "Then, What is Charisma? The Role of Audio-visual Prosody in L1 and L2 Political Speeches," *Proc. Phonetik & Phonologie*, Munich, Germany, pp. 1–3, 2016
- [28] Novák-Tót E., Niebuhr O., Chen A., "A gender bias in the acoustic-melodic features of charismatic speech?" *Proc. 18th Interspeech*, Stockholm, Sweden, pp. 2248–2252, 2017.
- [29] Berger S., Niebuhr O., Peters B., "Winning Over an Audience – A Perception-based Analysis of Prosodic Features of Charismatic Speech," *43rd DAGA*, Kiel, Germany, pp. 1454–1457, 2017.
- [30] Bosker H.R., "The role of temporal amplitude modulations in the political arena: Hillary Clinton vs. Donald Trump," *Proc. 18th Interspeech*, Stockholm, Sweden, pp.1–5, 2017.
- [31] van Stel A., Carree M., and Thurik R., "The Effect of Entrepreneurial Activity on National Economic Growth," *Small Business Economics* 24, pp. 311–321, 2005.
- [32] Delmar F., Shane S., "Legitimizing first: organizing activities and the survival of new ventures," *Journal of Business Venturing* 19, pp. 385–410, 2004.
- [33] Fisher G., Kuratko D., Bloodgood J.M., Hornsby J.S., "Legitimate to whom? The challenge of audience diversity and new venture legitimacy," *Journal of Business Venturing* 32, pp. 52–71, 2017.
- [34] Simpson, A.P., "The first and second harmonics should not be used to measure breathiness in male and female voices," *Journal of Phonetics* 40, 477-490, 2012.
- [35] Traunmüller, H., "Conventional, biological, and environmental factors in speech communication: A modulation theory," *Phonetica*, vol. 51, pp. 170–183, 1994.
- [36] Lindh, J. and Eriksson, A., "Robustness of long time measures of fundamental frequency," in: *Proceedings of Interspeech 2007*, pp. 2025–2028, 2007.
- [37] Boersma, P. and Kovacic, K., "Spectral characteristics of three styles of Croatian folk singing," *Journal of the Acoustical Society of America* 119, pp. 1805–1816, 2006.
- [38] Frøkjær-Jensen, B. and Prytz, S., "Registration of voice quality," *Briël & Kjaer Technical Review*, vol. 3, pp. 3–17, 1976.
- [39] Leino, T., "Long-term average spectrum in screening of voice quality in speech: Untrained male university students," *Journal of Voice*, vol. 23(6), pp. 671–676, 2009.
- [40] Guzman, M., Correa, S., Munoz, D. and Mayerhoff, R., "Influence on spectral energy distribution of emotional expression," *Journal of Voice*, vol. 27(1), 129.e1–129.e10, 2013.
- [41] Bele, I. V., "The speaker's formant," *Journal of Voice*, vol. 20(4), pp. 555–578, 2006.
- [42] Nolan, F. and Grigoras, C., "A case for formant analysis in forensic speaker identification," *International Journal of Speech, Language and the Law*, vol. 12, pp. 143–173, 2005.
- [43] Moos, A., "Long-term formant distribution as a measure of speaker characteristics in read and spontaneous speech," *The Phonetician*, vol. 101/102, pp. 7–24, 2012.
- [44] Bushman, B.J. and M.C. Wang, "A procedure for combining sample correlation coefficients and vote counts to obtain an estimate and a confidence interval for the population correlation coefficient," *Psychological Bulletin* 117, pp. 530-546, 1995.
- [45] Cabane, O.F., *The Charisma Myth: How Anyone Can Master the Art and Science of Personal Magnetism*. New York: Penguin, 2012.