Shapes and timing in charismatic speech - Evidence from sounds and melodies

Niebuhr, Oliver; Thumm, Jana; Michalsky, Jan

Published in:
Proceedings of the 9th International Conference on Speech Prosody 2018

DOI:
10.21437/SpeechProsody.2018-118

Publication date:
2018

Document version:
Final published version

Document license:
Unspecified

Citation for pulished version (APA):
Niebuhr, O., Thumm, J., & Michalsky, J. (2018). Shapes and timing in charismatic speech - Evidence from sounds and melodies. In *Proceedings of the 9th International Conference on Speech Prosody 2018* (pp. 582-586). International Speech Communication Association (ISCA). https://doi.org/10.21437/SpeechProsody.2018-118

9th International Conference on Speech Prosody 2018
13-16 June 2018, Poznań, Poland

Shapes and timing in charismatic speech - Evidence from sounds and melodies

Oliver Niebuhr[1], *Jana Thumm*[1], *Jan Michalsky*[2]

[1]SDU Electrical Engineering, Mads Clausen Institute, University of Southern Denmark
[2] Institute of German Studies, University of Oldenburg, Germany

oniebuhr@mci.sdu.dk, jana.thumm@gmx.de, j.michalsky@uol.de

Abstract

Our paper presents a phonetic analysis at the intersection of segments and prosodies. We look in detail at the previous finding that high pitch and a clear pronunciation contribute to a speaker's perceived charisma. To that end, we compare two popular CEOs, Steve Jobs and Mark Zuckerberg, who are known (from informal observations and formal perception experiments alike) to be more or less charismatic speakers, respectively. The results of our between-speaker comparison suggest that high pitch not only involves the F0 level but also the timing and shaping of pitch accents, and that a clear pronunciation not only refers to a large vowel space but also to the timing and shaping of consonant patterns in terms of fewer place assimilations and clearly separated voiced and voiceless stops. Perspectives for future research and implications for the training and evaluation of charismatic speech are discussed.

Index Terms: charisma, pitch accent, assimilation, F0.

1. Introduction

Some speakers draw us under their spell. We cannot help but listen to them, we believe in what they tell us, and we are willing to adopt their opinions, attitudes, and/or agendas. Attracting attention as well as gaining and persuading followers without having any formal authority is the essence of charisma [7]. The "wow effect" of charisma leads to more fruitful brainstorming sessions [1], results in better learning outcomes of students [2], helps raise more start-up funding [3], and makes a product or service appear more credible and likable to customers [4]; moreover, it can help people climb up the career ladder [5]. Previous studies also demonstrated that charisma is not a mysterious talent of a few gifted people, as was originally claimed by [6], but a continuously varying skill that anyone can learn and improve [7].

This learning and improving requires, though, that we understand the mechanisms that make a speaker sound charismatic. In particular, this applies to the mechanisms that underlie *how* people say something as these phonetic mechanisms are probably more important than lexical ones for a speaker's charismatic impact [8,9]. Thus, phonetics should actually be a key field of research on charisma. But, in reality, charisma plays a rather subordinate role in phonetics, and phonetic research on charisma is given insufficient consideration in other charisma-related disciplines. For example, the recent call from researchers of psychology and business to base speaker charisma on "unobtrusive and objective measures that do not rely on perceptions of raters" [10:308] resembles the one that was already made more than a decade ago by two phoneticians who stressed the need for "defining a set of objective measures of charisma" [11:1]. To overstate the case: Current research and practice in charismatic speech still rely primarily (i.e. outside the field of work of a few phonetic experts) on vague descriptive terms and imitation-based methods that were already used by Aristotle and Quintilian hundreds of years ago [12].

This is true despite the fact that phonetics has made significant steps towards putting speaker charisma on a new digital and quantitative footing. Phonetic researchers identified, with a focus on political speakers, a number of acoustic-prosodic features that correlate (in culture-specific ways) with speaker charisma, such as the level, range, and variability of F0 and intensity, speaking rate, and the utterance-to-pause duration ratio. Previous phonetic studies also broke down charisma into simpler adjectives with which the relevant acoustic-prosodic features are strongly correlated [13,14,15,16,17,18,19].

Recently, we showed that the acoustic-prosodic features of charisma that were identified in previous studies can be transferred from political to business speakers like Steve Jobs and Mark Zuckerberg, on whose data we also report in the present paper. Moreover, our research addressed the issue of gender-specific behavior (of both speakers and listeners) [26] and added further parameters like the frequency (cpm) of emphatic accents and disfluencies, voice quality (HNR, jitter, H1-H2, 1-5/5-8 kHz ratio), rhythmic variability (VarcoV, %V), and the singer's formant to the list of charisma-relevant features [20]. In addition, we extended the list of charisma-relevant features into the domain of sound segments, supporting the traditional claim of rhetoric that a "clear" and "crisp" articulation of "every phrase and word" [21:158] "is imperative to develop charisma" [22:138].

Yet, for being able to effectively assess, measure, and train charismatic speech, a more fine-grained understanding is needed of how the individual relevant parameters contribute to charisma. That is, after having set the parametric framework, phonetic research now needs to shift its focus gradually from the parameters themselves to their phonetic and phonological details. The present paper represents an initial step in this direction, addressing the acoustic differences between the more/less charismatic business speakers Steve Jobs (SJ) and Mark Zuckerberg (MZ), see Figure 1.

Figure 1: *Former CEO of Apple Steve Jobs (left) and current CEO of Facebook Mark Zuckerberg (right).*

There are three reasons why we use the SJ and MZ data in this study. First, we have already addressed this speaker pair in previous analyses, in the course of which we created a prosodically and segmentally fully annotated corpus of their

10.21437/SpeechProsody.2018-118

major keynotes. This corpus also provides the basis for the present study. Second, the public opinions about SJ's and MZ's speaker charisma are numerous and clearly different. While SJ is celebrated as "a master of the art of effective and persuasive speaking" [40], MZ's presentation skills were described to be "rough enough to impact Facebook's perception in a negative way" [41]. Third, these assessments from the media are consistent with the result of our own formal perception experiment, see, for example, [42]. The experiment was based on randomly selected and de-lexicalized and de-personalized (i.e. low-pass-filtered) 30-second excerpts of keynote speeches given by SJ and MZ. The excerpts were integrated in a larger set of de-lexicalized stimuli and played, in individually randomized orders, to 98 participants whose task was (i) to rate the speaker's charisma on a scale from 0-10, (ii) to estimate the speaker's leadership/management experience (0-10 years), (iii) to put a figure on the likelihood of investing own money into the speakers company (0-100%). As is shown in Figure 2, SJ clearly and statis-tically significantly outperformed MZ on the charisma question ($t[97]=11.5$, $p<0.001$) and the management/ leadership question ($t[97]=24.8$, $p<0.001$). The investment question yielded no differences between the two speakers.

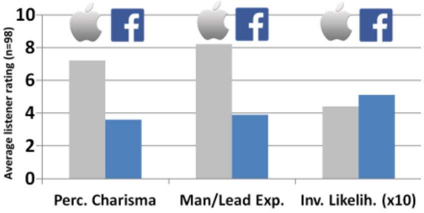

Figure 2: *Charisma and charisma-related judgments made by 98 listeners for de-lexicalized and de-personalized stimuli excerpted from keynote speeches of SJ and MZ.*

2. Questions and expectations

One consistent finding in previous studies is that more charismatic speech is characterized by higher F0 values [13-17]. In this paper, we investigate this F0 raising more closely by looking at prosodic details of pitch accents. The question is whether the higher F0 of more charismatic speakers manifests itself only as a holistic increase in F0 level (if at all, see [42]), or whether being a higher-pitched more charismatic speaker additionally involves using more H* pitch accents and shaping these accents such that they enhance high F0 values, for example, through F0 slopes that are longer and more convex, i.e. slopes that rise faster to a higher F0 level. If this is true, then we would expect the more charismatic speaker in our speaker pair, i.e. SJ, (i) to not just to show a generally higher F0 level than MZ but also to produce more H* pitch accents and (ii) to enhance the high F0 values underneath the pitch-accent peaks through changes in peak shape more strongly than MZ.

In addition, we further examine the traditional claim of rhetoric that a clear articulation adds to speaker charisma. This claim is, so far, only supported by evidence from contrastive vowel-space analyses of SJ and MZ. More specifically, SJ was found to use a significantly larger vowel space than MZ (along both the F1 and the F2 dimension, [42]). In this paper, we investigate this pronunciation-related segmental parameter in more detail by looking at stop-consonant voicing and place assimilation of alveolar consonants. The question is whether what applies to vowels also applies to consonants. That is, if more charismatic speech means clearer speech also for consonants, then we would -- in view of the enhanced-segmentability finding of [43] -- expect SJ (i) to produce greater phonetic contrasts (in terms of duration and voicing) between his voiced and voiceless stops and (ii) to show fewer assimilations of /t/, /d/, and /n/ to other places of articulation than MZ.

3. Method

Analogous to the studies of [11,17], our analysis was based on sections of official and thus strongly conventionalized speeches of SJ and MZ. Moreover, taking into account Conger's model of the complexity and contextual embedding of charisma [27], we narrowed down the selection of speeches to a single genre. The genre we used are product presentations, also because these globally broadcasted introductions of new products are particularly often referred to in the literature when it comes to speaker charisma [28].

For SJ, we used two of his most well-known and influential keynotes: the presentation of the iPhone 4 in 2010, and the presentation of the iPad 2 in 2011. Twenty-two (10+12) minutes of speech were extracted from the two presentations, which corresponds to about 12,000 individual speech sounds and 692 prosodic phrases. MZ's data was extracted from two of his recent keynotes (held in 2014, 2015) at Facebook's annual F8 meeting. The structures of MZ's selected keynotes are similar to those of SJ: Showcasing the new product (new website features or functionalities in the case of Facebook) is followed by notes on the company's last year's achievements, position in the market, and future growth strategy. MZ's keynote excerpts also comprise about 22 (11+11) minutes of speech, which corresponds to about 13,700 speech sounds and 536 prosodic phrases. The two speech samples of MZ and SJ were extracted from the middle of their keynotes to exclude the usual opening and closing rituals of their presentations and potential biases due to the speakers' warm-up phases.

We considered 22 minutes a sufficiently representative and rich sample of a speaker's speech, not least because the number of prosodic phrases contained in 22 minutes of speech (500-700) is about 2-3 times as high as the number of sentences elicited for most read-speech analyses on F0-peak alignment, coarticulation, phonetic reduction, etc.; and 22 minutes is also more than what is commonly needed for forensic speaker recognition and related phonetic analyses.

The audio files of SJ and MZ were obtained from high-quality YouTube videos of the corresponding keynote presentations. Video and audio were separated, and the audio files were saved in the uncompressed WAV file format.

The SJ-MZ speech corpus was prosodically and segmentally annotated. The prosodic annotation was conducted on an auditory basis by a trained annotator (JT) who used the AM-related DIMA framework [30]. For the segmental annotation, the speech samples of SJ and MZ were first orthographically transcribed by an English speaker and then submitted to automatic segmentation based on WebMaus and Darla [31,32]. Based on the annotated corpus, the following sets of measurements were taken manually by the authors.

Pitch-accent related measurements (including only nuclear and prenuclear accents with two or more unaccented syllables on both sides)

- Number of H* accents, cf. [37]
- Number of HL* and L*H accents, cf. [37]
- For all three types of accents: Duration of the rising and falling slope (in ms)

- Shape of the rising and falling slope in terms of a variant of the range proportion measure RP developed by [33]; RP is the F0 range of the first half of the rising or falling slope of a pitch-accent peak divided by the total F0 range of that slope (values >0.5 indicate convexity, values <0.5 indicate concavity)

Pronunciation-related segmental measurements

- Separately for stressed and unstressed syllables: Total closure duration of syllable-initial /ptk/ and /bdg/ stop consonants followed by vowels or sonorants (in ms)
- Number of /bdg/ stops realized mainly (>50%) voiced (determined from the signal elongation patterns in the displayed waveform during stop closure)
- Number of place assimilations (in content words) of alveolar /tdn/ to labial or velar consonants of following syllables (determined on an auditory basis)

Each measurement yielded a pair of independent samples (SJ vs MZ). The statistical treatment of these independent samples was based on conservative tests. We used the z-score proportion tests for counts and the Welch t-test for acoustic data. Compared to the regular Student's t-tests for independent samples, the Welch t-test performs better for heterogeneous variances and for unequal samples sizes [34], both of which applied to parts of our data.

As for the samples sizes, note that the measurements are based on a random selection of at least about 50 tokens for stop consonant durations, at least 100 tokens for all other segmental features and about 200 tokens for all prosodic features. A random selection was necessary in order to reduce the workload of the 2nd author who conducted the analyses as part of an international student internship program.

4. Results

Are there differences in pitch-accent type (i.e. timing) and shape between SJ and MZ? The results addressing this question are summarized in Figures 3(a)-(b). They show two things. Firstly, SJ uses more H* pitch accents in his speech than MZ. This difference is significant according to a z-score proportion test (z=2.4, p<0.05). Secondly, the peak-shape patterns of SJ strongly enhance high F0 underneath the peak. Compared to MZ, SJ's pitch-accent peaks are characterized by longer rises (t=2.11, p<0.05) as well as by rising and falling slopes that are more convex than those of MZ (t=2.29, p<0.05; t=2.65, p<0.01). Note that this is true independently of the type of pitch accent (H*, HL* or L*H), although there are also pitch-accent specific slope durations and shapes that we cannot report here in detail, but which are in line with [44].

Regarding the overall nature of the F0 differences between SJ and MZ, note that the rising and falling slopes of MZ's pitch accents are also convexly shaped, like those of SJ; and, like for SJ, MZ's mean F0 level is also raised compared to that of the typical male mid-age American English speaker (120 Hz, see [36]). However, the convexity in MZ's pitch accents is weaker, and his F0 raising not as strong as that of SJ. Yet, in combination, our results suggest that the difference between SJ and MZ is not a qualitative one (in the sense that SJ does something that MZ does not) but a quantitative one (in the sense that MZ does something that SJ does in more a pronounced or skillful way).

The results of the acoustic stop-consonant measurements are summarized in Figures 4(a)-(b). With respect to the phonetic differences between /ptk/ and /bdg/, Figure 4(a) shows the following: While both speakers produce /ptk/ longer than /bdg/ (t=4.55, p<0.001) and stops in accented syllables longer than in unaccented syllables (t=3.84, p<0.001), the stop durations increase significantly more from /bdg/ to /ptk/ for SJ than for MZ, cp. in Figure4(a) how much longer the dark bars of /ptk/ are in relation to the light bars of /bdg/. This stronger increase in stop duration from /bdg/ to /ptk/ for SJ applies in particular to the /b/-vs-/p/ and /d/-vs/-t/ differences in unaccented syllables (t=1.98, p<0.05; t=2.83, p<0.01) and to the /g/-vs-/k/ difference in accented syllables (t=2.61, p<0.01).

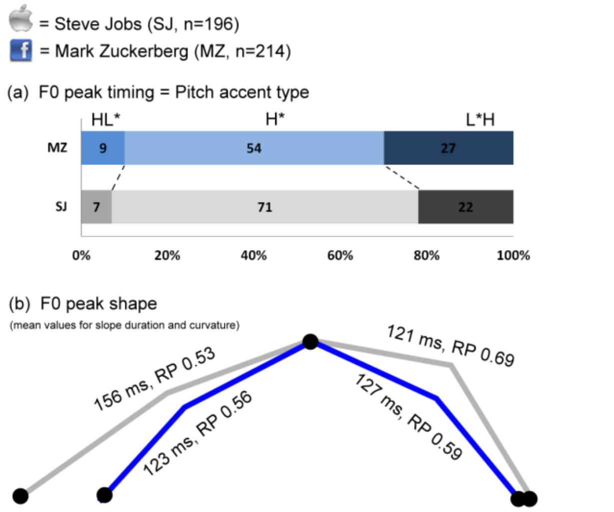

Figure 3: *(a) Distributions of pitch-accent types; (b) mean values for range proportion (RP) and slope duration.*

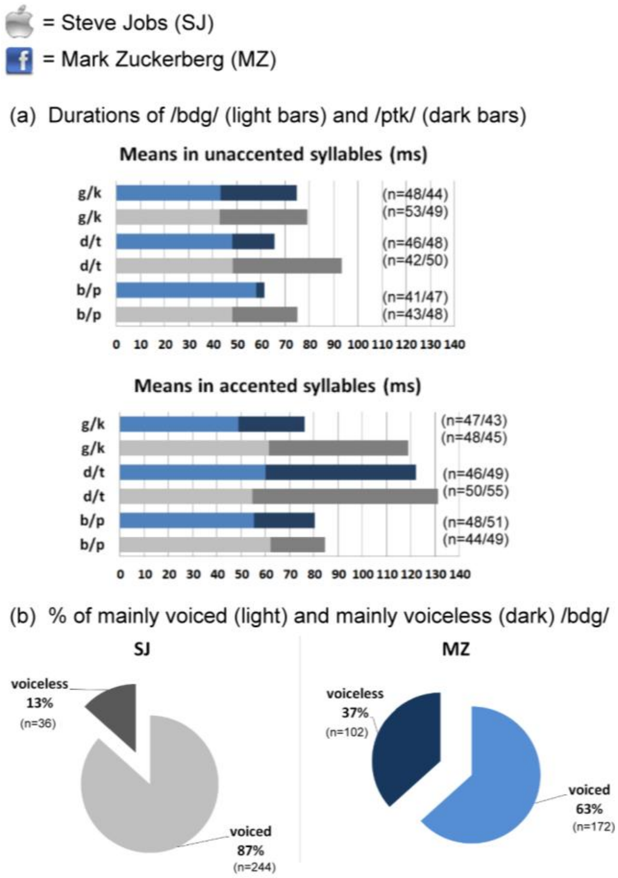

Figure 4: *(a) Mean stop-consonant durations; (b) percentages of SJ and MZ producing mainly voiced and voiceless /bdg/.*

Moreover, Figure 4(b) shows that the percentage of phonologically voiced stops (/bdg/) that are actually realized mainly voiced is higher for SJ than for MZ. This difference is also significant for both unaccented ($z=3.38$, $p<0.001$) and unaccented syllables ($z=2.72$, $p<0.01$), and it occurs although SJ's /bdg/ realizations are not shorter (and hence more likely to be inherently voiced) than those of MZ, see Figure 4(a).

Finally, Figure 5 shows that, according to our auditory analysis, the majority of MZ's syllable-final alveolar /tdn/ consonants was realized assimilated to the place of articulation of the labial or velar consonant in the following syllable. In contrast, for SJ, this place assimilation only occurs for a minority of syllable-final alveolar /tdn/ consonants, which represents a significant difference to MZ in a z-score proportion test ($z= 2.14$, $p<0.05$). Note that this difference is similar independently of whether or not we look separately at assimilation frequencies within and across (content) word boundaries.

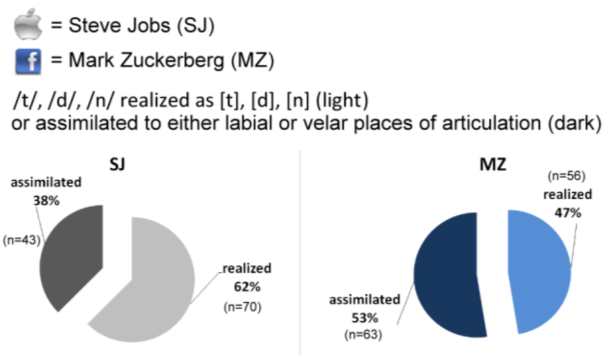

Figure 5: *Anticipatory place assimilations (in %) of /t/, /d/, and /n/ at syllable boundaries.*

5. Discussion

The present paper was about shape and timing in charismatic speech. We looked at shape in the sense of the duration and curvature of pitch-accent slopes and the articulatory shaping of pairs of sound segments, which then does or does not lead to anticipatory alveolar place assimilation across syllable and word boundaries. With respect to timing, the paper was concerned with the duration of stop consonants, the co-ordination of voicing with stop closures, and with timing in the form of pitch-accent categories whose differences manifest themselves primarily in how a F0 peak pattern is aligned relative to the accented syllable [37,44]. Based on statements in the literature that (i) a higher F0 level and (ii) a clearer (i.e. more hyperarticulated) pronunciation make a speaker sound more charismatic, we investigated with respect to (i) whether the higher F0 level also involves more H* accents and, in general, pitch accents with longer and more convexly shaped slopes, and whether with respect to (ii) a clearer articulation also involves fewer alveolar place assimilations and greater phonetic distances between voiced and voiceless stop consonants. Our investigation was based on comparing two popular business speakers, a more charismatic one (Steve Jobs) and a less charismatic one (Mark Zuckerberg). In line with our expectations derived from the speakers' different charisma levels, we found that SJ used more H* accents than MZ and produced pitch accents with slower rising and overall more convexly shaped slopes. SJ's speech was, compared to that of MZ, also characterized by fewer place assimilations and greater phonetic distances between /bdg/ and /ptk/ in terms of durational differences and the presence of voicing inside /bdg/.

What we can conclude from these findings is that, for prosody, it is worth looking in more detail at each of the acoustic parameters that are known from previous phonetic studies to be correlated with speaker charisma [13-19,25-26]. Probably each parameter can be broken down into a set of prosodic components. Performing these in-depth analyses is important for assessing and teaching charismatic speech. For example, simply knowing as a charisma teacher or learner that 'higher' (pitched) means 'better' is not sufficient if 'higher' is not restricted to the F0 level alone (if at all, see [42)], but also includes pitch-accent types (timing) and shapes. Furthermore, it becomes obvious in this application-oriented context that follow-up studies have to scrutinize how much the accent types and shapes actually contribute to the acoustical and perceptual increase in a speaker's average pitch level, and, on this basis, what the effectiveness hierarchy of the individual pitch-related components is within the total effectiveness of the charisma factor 'higher'. Analogues questions arise for all other prosodic parameters.

For the level of sound segments, we can conclude from our study that articulatory patterns, analyzed in terms of the hypo-hyper scale [38,42], have to become an integral part of phonetic research on speaker charisma. Our findings in combination with the constant call of rhetoric for a clear articulation (e.g, [21,22]) point to a general systematic difference between more and less charismatic speakers that affects many (if not all) segmental features and deserves a detailed and objective phonetic description. Follow-up studies in this area should also investigate whether there is an upper threshold up to which a clearer pronunciation is more charismatic, as is suggested by the results of a recent perception experiment [42]. Such a threshold probably differs as a function of speaker gender [39], speaking style, and the situational context in general [38]. For example, listeners might pose higher demands on clarity in the pronunciation of female speakers, but might find an all too clear pronunciation in a meeting with a small group of colleagues less charismatic than in front large audience while giving a business presentation.

Another relevant threshold factor is probably speaking rate. This parameter is, like a clear pronunciation, positively correlated with speaker charisma. However, the two parameters mutually limit each other as to the amount they can be increased. That is, there is an upper limit until which clear and fast are still biomechanically compatible. At this point the question arises what is more important, a high speaking rate or a clear articulation. MZ is a faster speaker than SJ. But, SJ is perceived to be more charismatic, and he obviously outperforms MZ in many aspects of articulatory clarity, some of which are probably related to a lower speaking rate (like the lower assimilation frequency). Thus, our working hypothesis for a follow-up study is that a striving for clearer articulation is more important than striving for a higher speaking rate.

The speaking rate example further shows that the interplay of segmental and prosodic factors in the production and perception of charismatic speech also needs to be addressed by phonetic studies. That is, it needs to be tested to what degree a prosodic key concept of charisma, i.e. variability, also extends into the domain of sound segments in the form of an alternation between higher and lower levels of speech reduction and to what degree this alternation is (or should be) identical to the one between content and function words.

6. References

[1] Pentland A., *Honest Signals - How they shape our world*. Cambridge: MIT Press, 2008.

[2] Towler A.J., “Effects of charismatic influence training on attitudes, behavior, and performance,” *Personnel Psychology 56*, pp. 363-381, 2003.

[3] 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.

[4] 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.

[5] Bodow S., “Charmed I'm sure,” *USA Today*, 2012.

[6] Weber, M., *The Theory of Social and Economic Organization*. New York: Free Press, 1947.

[7] Antonakis J., Fenley M., Liechti S., “Can charisma be taught? Tests of two interventions,” *Acad. Manag. Learn. Educ. 10*, pp. 374–396, 2011.

[8] 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.

[9] Awamleh R., Gardner, W.L., "Perceptions of leader charisma and effectiveness: The effects of vision content, delivery, and organizational performance. *The Leadership Quarterly 10*, pp. 345–373, 1999.

[10] 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.

[11] Rosenberg A. and Hirschberg J., “Acoustic/Prosodic and lexical correlates of charismatic speech,” *Proc. Eurospeech 2005, Lisbon, Portugal*, pp. 1–4, 2005.

[12] Niebuhr O., Tegtmeier 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.

[13] Touati P., “Prosodic aspects of political rhetoric,“ *Proc. ESCA Workshop on Prosody, Lund, Sweden*, pp. 168–171, 1993.

[14] 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.

[15] 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.

[16] 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.

[17] Rosenberg A., Hirschberg J., ”Charisma perception from text and speech,” *Speech Communication 51*, pp. 640–655, 2009.

[18] 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

[19] 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.

[20] Niebuhr O., "On the phonetics of intensifying emphasis," *Phonetica 67*, pp. 170-198, 2010.

[21] K.W. Mortensen, *The Laws of Charisma: How to Captivate, Inspire, and Influence for Maximum Success*. New York: Amacom, 2011.

[22] R. Camper Bull, *Moving from Project Management to Project Leadership: A practical guide to leading groups*. Boca Raton: CRC, 2010.

[25] Berger S., Niebuhr O., Peters B., “Winning Over an Audience – A Perception-based Analysis of Prosodic Features of Charismatic Speech,” *Proc. 43rd Annual Conference of the German Acoustical Society, Kiel, Germany*, pp. 1454–1457, 2017.

[26] 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.

[27] Conger J.A., Kanungo R.N., *Charismatic Leadership in Organizations*. London: SAGE, 1998.

[28] Sørensen, L.S., *How to Grow an Apple: Did Steve Jobs Speak Apple to Success*. MA thesis, Aarlborg University, Denmark, 2013.

[30] Kügler F., Smolibocki B., Arnold D., Braun B., Baumann S., Grice M., …Wagner P., "DIMA – Annotation guidelines for German intonation," *Proc. 18th International Congress of Phonetic Sciences, Glasgow, UK*, pp. 1–5, 2015.

[31] Strunk J., Schiel F., Seifart, F., "Untrained forced alignment of transcriptions and audio for language documentation corpora using WebMaus," *Proc. 9th International Conference on Language Resources and Evaluation, Reykjavik, Iceland*, pp. 3940-3947, 2014.

[32] Reddy S., Stanford J., "Toward completely automated vowel extraction: Introducing DARLA," Linguistics Vanguard, 2015.

[33] Dombrowski E., Niebuhr, O., “Acoustic patterns and communicative functions of phrase-final F0 rises in German: Activating and restricting contours,” *Phonetica 62*, pp. 176-195, 2005.

[34] Rietveld T., van Hout R., "The t test and beyond: Recommendations for testing the central tendencies of two independent samples in research on speech, language and hearing pathology," *Journal of Communication Disorders 58*, pp. 158-168, 2015.

[35] A.P. Simpson, “Phonetic differences between male and female speech,” *Lang. Linguist. Compass 3*, pp. 621–640, 2009.

[36] Pépiot, E. "Male and female speech: a study of mean f0, f0 range, phonation type and speech rate in Parisian French and American English speakers," *Proc. 7th International Conference of Speech Prosody, Dublin, Ireland*, pp.305-309, 2014.

[37] Kleber F., "Form and function of falling pitch contours in English," *Proc. 3rd International Conference of Speech Prosody, Dresden, Germany*, pp. 61-64, 2006.

[38] Lindblom B, "Explaining phonetic variation: A sketch of the H & H theory," in W.J. Hardcastle, A. Marchal (eds), *Speech Production and Perception* (pp. 403-439), Dordrecht: Kluwer.

[39] Weirich M., Fuchs S., Simpson A., Winkler R., Perrier P., "Mumbling: Macho or Morphology?" *J. Speech Lang Hear Res 59*, pp. S1587-S1595, 2016.

[40] Heracleous L., "Why Steve Jobs is such a charismatic leader," *Warwick Business School, http://www.wbs.ac.uk/news/why-steve-jobs-was-such-a-charismatic-leader1/*, 2016.

[41] Gruener W., "On stage, Mark Zuckerberg is no Steve Jobs," *Tom's Guide, http://www.tomsguide.com/us/facebook-skype-video-calling-social-networking,news-11808.html*, 2011.

[42] Mixdorff H., Niebuhr O, and Hönemann A, "Model-based prosodic analysis of charismatic speech," *Proc. 9th International Conference of Speech Prosody, Poznan, Poland*, 1-5, 2018.

[43] Smijlanic R., Bradlow A.R., " Temporal organization of English clear and conversational speech," *J. Ac. Soc. Am 124*, pp. 3171-3182, 2008.

[44] Niebuhr O., "The signaling of German rising-falling intonation categories: The interplay of synchronization, shape, and height," *Phonetica 64*, pp. 174–191, 2007.