Speaker Charisma Analyzed through the Cultural Lens

Gutnyk, Anna; Niebuhr, Oliver; Gu, Wentao

Published in:
2021 12th International Symposium on Chinese Spoken Language Processing (ISCSLP)

DOI:
10.1109/ISCSLP49672.2021.9362100

Publication date:
2021

Document version:
Accepted manuscript

Citation for published version (APA):

Go to publication entry in University of Southern Denmark's Research Portal

Terms of use
This work is brought to you by the University of Southern Denmark. Unless otherwise specified it has been shared according to the terms for self-archiving. If no other license is stated, these terms apply:

- You may download this work for personal use only.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying this open access version.

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim. Please direct all enquiries to puresupport@bib.sdu.dk
Speaker charisma analyzed through the cultural lens

Anna Gutnyk²*, Oliver Niebuhr¹, Wentao Gu²

¹Centre for Industrial Electronics, University of Southern Denmark, Sønderborg
²Institute of Linguistics and Applied Linguistics, Nanjing Normal University, China
*gutnyk@163.com

Abstract
Speaker charisma is conveyed through multiple aspects: ideas, visions, and perceivable verbal and non-verbal behaviors. Among these perceivable behaviors, probably the most prominent are the acoustic features of one’s voice. We present here a cross-cultural study on charisma. We conducted a combination of acoustic-prosodic analysis and perception experiment based on speeches given by presenters from 6 different countries on 4 continents to shed initial light on the cultural differences in producing and perceiving charisma, as well as on the impact of speaker gender. Results show that charisma production and perception is affected by prosodic factors to different extend across both countries and speaker gender.

Index Terms: charisma, gender, culture, prosody, speech.

1. Introduction
East is not West. Cultures differ, and minds, feelings, and intentions in different societies intermesh in various ways [1]. A ‘charismatic template’ of one’s self is being influenced by culture, gender stereotypes and rhetoric tradition acquired in the society. Thus, the concept of a charismatic speaker cannot be the same for all countries, just like standards of physical attractiveness or delicious food around the globe.

Perceived speaker charisma is a complex, multi-component phenomenon that requires an interdisciplinary approach. So far, scientific studies have poorly addressed this multi-disciplinary complexity. Besides linguistics and phonetics, understanding charisma involves social sciences, political sciences, business sciences, psychology (social, personality and organizational), as well as ethology, biology, aesthetics, media science, physics, and, last but not least, pedagogy. A broad and profound understanding of speaker charisma can only emerge from close collaborations between these research disciplines.

2. Concept of charisma
Charisma has been a favored topic of discussion and research when it comes to defining, training, and practicing leadership. As Antonakis et al. [1] define it, charisma is a device for emotion-laden, values-based, symbolic leadership signaling. Phonetic research successfully filled the notion of charisma with acoustic-prosodic substance, thus making it a measurable and trainable matter [2], [3], [4], [5], [6], [7], [10].

People’s desire to become successful, to learn how to lead masses and how to persuade and charm people inspired phoneticians to develop tools and techniques for training charismatic speech [8], [9], [10]. As [11] point out: “despite the technological changes occurring in places of work, skills in oral communication [...] are most essential for career improvement” (p. 412). Moreover, charismatic speaking skills can be helpful not only in working life and career development, but also in everyday life: dating, teaching, negotiating prices, etc. Therefore, it is of vital importance to examine exponents of charisma and their application in different contexts.

Recent studies by [12] and [13] showed that the charismatic voice parameters found in politics also hold for business contexts, see also [14], [5], [3], [15], [16]. However, other aspects of charisma, such as cultural differences, universality of charisma – a direction in which [5] have taken the first steps – or speaker gender are yet to be explored.

The present study’s primary goal is to investigate whether charismatic speech has the same acoustic exponents in Chinese (CHIN), Ukrainian (UKR), Spanish (ESP), German (GER), Turkish (TURK), and Brazilian (BRAZ) cultures; also, we check if women display different charismatic characteristics than their male counterparts, and we investigate if there is a gender bias in the evaluation of charismatic speech [17], [18], [19], [20], [21], [22].

3. Acoustic features of charismatic speech
Phonetic research successively identified the acoustic parameters of charismatic speech and determined how and to what extent they contribute to perceived charisma. Charismatic speakers employ acoustic characteristics, such as an elevated rather than a lowered fundamental frequency (F0), an enhanced variation in F0, higher levels of vocal effort and intensity [23], [2], [14], [3], [5], [24], [25], [26], an increased speaking rate, a larger number of silent pauses within and between sentences, fewer and shorter filled pauses [23], [2], [3], [5], [12], [26], [27], a strategic placement of empathic accents [6], [27] as well as a higher level of articulatory precision [28], a greater acoustic-energy dynamics [29] and a more balanced spectral-energy distribution [26]. Lastly, more charismatic speakers partition their speech into smaller pieces of information by using more and shorter prosodic phrases [12] whose durations respect the auditory short-term memory [30].

4. Research review and hypotheses

4.1. Charisma and culture (H1)
Cullen et al. [31] crowd-sourced charisma ratings on an Irish politician’s speech and built automatic systems to detect charisma. By comparing the perception of charismatic cues by American and Irish listeners they discovered that a cultural divide exists between raters.

Biadsy et al. [14] compared experiments in which US American, Palestinian and Swedish subjects rated political speeches of Arabic speakers with respect to charisma. Results
showed that US Americans rated those Arabic speakers more charismatic who employed a faster speaking rate and spoke louder, but, at the same time, with a more varied/dynamic loudness pattern. Palestinians, by contrast, showed less sensitivity to these qualities. Swedish subjects rated those Arabic speeches more charismatic that had a higher minimum pitch and a lower pitch standard deviation. Thus, compared to US listeners, Swedish listeners considered higher pitched speech with a relatively compressed pitch range more charismatic.

Also examining political speech, Signorelli et al. [4] used speeches by Umberto Bossi, the leader of an Italian party, before and after a stroke. Results revealed that, for French listeners, the prosody in Bossi’s pre-stroke speech sounded more charismatic (the pitch contour in particular). Italian participants, on the other hand, perceived Bossi’s prosody to be more charismatic in the post-stroke condition.

D’Errico et al. [5] conducted a study that tested the effect of pitch contour, duration, intensity and voice quality on the perception of charisma traits in participants from two cultures, using a multi-dimensional adjective-based scale of charisma perception. Italians and French listeners differed in the interpretation of the “human sociability” (Calm-Benevolent) dimension of charisma. Italians associated it with a more high pitched voice and shorter pauses, whereas French listeners associated it with a low-pitched voice and long pauses.

In summary, the results of the above studies reveal a good deal of cross-cultural differences as to the perception of speaker charisma, related to pitch contour, duration, intensity and voice quality. On the assumption that the perceptual evaluation of charisma reflects the listeners expectations about how charismatic speakers should sound like in their respective countries, we put forward the following hypothesis.

H1. CHIN, UKR, TURK, ESP, GER, BRAZ speakers will produce different acoustic-prosodic patterns when they are giving a charismatic speech.

4.2. Charisma and speaker gender (H2a&b)

Another hardly addressed aspect of charisma is a potential gender bias. Females could sound less charismatic than males [25], [32]. That is, even all else equal (prosodic parameters and verbal content), male presenters seem to have an inherent advantage over female presenters [33], [26]. Management and entrepreneurship research discuss for a long time the so-called “gender gap”. It includes the fact that female speakers are less persuasive in front of potential investors even if they present to them, under controlled experimental conditions, the same slide deck in the same words and with a similar delivery [34], [35]. For example, the study by Brooks et al. revealed a significant effect of entrepreneurs’ gender on pitch success (odds ratio of 1.57, p < 0.01). Male entrepreneurs were 60% more likely to be successful with their investor ‘pitch’ than female entrepreneurs (note that ‘pitch’ is a technical term in business for a concise product- or idea-oriented speech). There was a significant interaction between gender and attractiveness on the likelihood of pitch success (p < 0.05). Both professional investors and nonprofessional evaluators preferred pitches presented by male entrepreneurs compared to pitches held by female entrepreneurs, even if the content of the pitch was exactly the same.

Niebuhr et al. [10] demonstrated that, even when being judged by an automatic charisma scoring system called PASCAL, female speakers get significantly lower prosodic charisma scores than males. Although, women benefit more from prosodic charisma training then men, PASCAL measurements at the end of the course showed that they women did not surpass men’s scores, only approximately reached their level.

In another experiment, Niebuhr et al. [33] compared three famous charismatic speakers: two females (Oprah Winfrey, Ginni Rometty) and one male (Steve Jobs). The comparison included four prosodic features (i.e. hesitations, emphatic accentuation, speaking rate, and acoustic energy level). The two female speakers were judged by listeners to be similarly charismatic as the male speaker; but, in order to be at this eye level with the male speaker, they actually had to be better in their prosodic features than the male speaker.

Based on these largely consistent findings across research disciplines and methods, and with reference to how the countries involved in our speaker sample are ranked in the UN gender-gap report of the World Economic Forum (WEF) from 2017, we put forward the following two hypotheses.

H2a. Acoustic-prosodic features of female speaker from the society with a high gender inequality index (WEF report 2017: TURK #130, CHIN #106, BRAZ #92, UKR #59) will reveal stronger difference between male and female speech than in the countries with lower index (ESP #8, GER #10).

H2b. Gender gap will benefit male speakers in pitch evaluation.

4.3. Charisma and rhetoric tradition (H3)

A leader will be perceived as more charismatic if she expresses him/herself coherently with cultural expectations. According to Geert Hofstede, culture is ‘the collective programming of the mind’. And although all cultural goods undergo changes caused by cultural diffusion [36], Xing Lu [37] argues that “a new understanding of rhetoric” should start with “an understanding of the realm, role, and function of rhetorical concepts derived from and addressed to cultural forces and social contexts”.

Rhetoric means are one of the main components of charisma. Rhetoric was first conceived by Aristotle as the “available means of persuasion”. These means were codified into canons in classical Rome and have since then been a central part of Western education to train speakers, writers and politicians to effectively move their audiences. Speakers try to convince the audience to do some action by exploiting the three strategies postulated by Aristotle: Logos (the rational argument), Pathos (the appeal to the audience’s emotions), and Ethos (the character of the speaker), see also Poggy [42].

However, according to the relativity of rhetorical theories, the standards of rhetoric in the West are not universal. They are only expressions of Western culture, applicable within the context of Western cultural values [38]. In other words, rhetorical standards are more or less determined and affected by specific cultural traits. Rhetoric is intertwined with, and inseparable from philosophy, religion, ethics, psychology, politics, and social relations. The heritage of Western rhetoric owes a great deal to the doctrines of Aristotle and Cicero. Similarly, the heritage of Chinese rhetoric is heavily indebted to the strands of Buddhism, Confucianism, and Taoism which, for centuries, have run through the Chinese culture. For example, argument, as a means of persuasion, serves as the essential constituent in Western rhetoric. In Chinese rhetoric however, argument—like the concept of eloquence—is heavily deprecated, for “it is equated with contentiousness,
with exaggerating differences, with decreasing mutual understanding, with undermining harmony” [39]. A further point of contrast between early Chinese and Greek rhetoric was that the ability to speak well and persuasively in public was essential to the ambitious Athenians of the fifth and fourth centuries BC. That is, people were expected to participate in politics. In contrast, public speaking of this sort has had little place in Chinese political life [43]. Even nowadays, whenever someone is invited to speak to the public, Chinese speakers remember their traditional saying that “illness finds its way in through the mouth and disaster finds its way out through the mouth” [44]. Thus, China has a much stronger written than oral rhetorical tradition.

Muslims, on the contrary, have a very strong oral tradition. In fact, prophets used charisma to communicate the vison of Allah and attract religious followers. Allah’s prophet was renowned for his eloquence and clarity of speech. “Listener involvement is heightened, and the aesthetics of style and audience relations may supersede the informational aspects of a message” [45]. Listeners noted that religious leaders spoke simple language, organized in short sentences and many silent pauses. These rhetorical strategies are viewed as a means for retaining the audience’s attention as well as making the message agreeable so as to build rapport [45].

In a nutshell, rhetoric tradition is one of the main forces influencing stereotypes of how a charismatic speaker should sound. As cultures and women’s positions differ in each society, so do rhetorical preferences around the globe. We consider ESP, GER, UKR and BRAZ as following the Western rhetoric tradition, TURK following the Muslim rhetoric tradition, and CHIN as following the Confucian rhetoric tradition. We predict that charismatic cues of the speakers will reflect these individual rhetorical fingerprints.

H3. Charismatic speech of a specific country will reflect a particular pattern of rhetoric tradition.

5. Production study

5.1. Participants and materials

The countries/cultures/languages that are currently involved in this project are: Mandarin Chinese, German, Brazilian Portuguese, Ukrainian, Spanish, and Turkish. All languages were represented by native speakers who where, moreover, selected to be similar in terms of their education level, age and public-speaking experience. Each country is represented by 10 male and 10 female speakers. Since 6 countries are involved, the total speech corpus recorded includes 6 x 20 = 120 speakers. For the recordings we used a prepared sales pitch for a newly developed smart phone application that tracks employees’ work time. This sales pitch was the same for all 120 speakers, i.e. an award-winning pitch taken from the e-learning course on “How to write a killer elevator pitch”. The sales pitch was translated by professional interpreters into the 6 languages. The interpreters were native speakers of the target language and used an English version of the text for their translation.

Our speakers received a written text of the product presentation one day prior of being recording, in order for them to get acquainted with the text. In the recording session, speakers were asked to perform the product presentation two times and as charismatically as possible: once addressing an imagined male and once addressing an imagined female audience (of potential investors). Performing the sales pitch took between 60-90 seconds.

5.2. Method

Phonetic researchers identified, with a focus on political speakers, a number of acoustic-prosodic features that correlate with speaker charisma. We took those features as the basis for our study. Thus, the acoustic analysis included the 9 prosodic parameters below. Measurements of all parameters were conducted automatically by means of PRAAT scripts and checked manually for outliers. Note that these checks also removed critical micro-prosodic perturbations from the results data (critical insofar as to they caused measurement errors, e.g., with respect to pitch range, pitch variability or tempo). All other micro-prosodic perturbations remained in the data, as we assumed that, across 120 speakers and about 30 minutes of speech per language, they would turn into statistical noise.

- Pitch level (F0 median based on 90th percentile),
- Pitch range (num. of octaves, 90th percentile),
- Pitch variability (F0 standard deviation),
- Tempo (speaking rate in syll./s, excluding pauses),
- Number of silent pauses (> 200ms),
- Utterance duration (s),
- Voice-quality jitter (ppq5),
- Voice-quality shimmer (ppq5),
- Harmonics-to-noise ratio (HNR, dB).

5.3. Results

For assessing the statistical significance of the results, the result of study 1 was analyzed in a MANOVA. Based on that, we found that the pitch level (F0 median) and the pitch variability (F0 standard deviation) were highest for GER and lowest for ESP speakers (p<0.001). Other countries showed intermediate levels in these parameters. CHIN subjects gave the sales pitch at the lowest tempo – BRAZ and UKR at the highest tempo (p<0.01). Furthermore, HNR values were largest (i.e. voices were least breathy) for CHIN, and smallest (i.e. most breathy) for UKR and TURK speakers (p<0.05).

Lastly, compared to GER and ESP speakers, both UKR and ESP speakers inserted significantly more silent pauses in the given text, hence speaking in shorter prosodic phrases (p<0.001). Figures 1 illustrates the results for four parameters which we assume to be of most general interest.

Thus, our acoustic evidence supports H1 and, in some way, also H3. Indeed, speakers from different languages and cultures differ in their production of prosodic exponents of perceived speaker charisma (elicited through a sales pitch).

Our speech material from the 6 compared languages showed additionally that women spoke at a higher pitch level than men, independently of country (p<0.001). The same applied to pitch variability (p<0.001). Moreover, women's tempo was slower than that of men, and their voices were also characterized by lower jitter and HNR values (p<0.01 for both parameters). That is, female voice qualities were breathier but less "shaky" or irregular in terms of pitch. Finally, women also made fewer pauses in their sales pitch presentations than men (p<0.05). The speakers from UKR, TURK and BRAZ differ more strongly with respect to gender prosody than the speakers from GER, ESP and CHIN (p<0.05). This finding suggests a positive answer to our H2a.
6. Perception study

6.1. Method

The perception experiment was to test, if women are rated as charismatic as men (all else equal), or if they suffer from a gender-gap discrimination. To ensure that the content of the speech is properly understood, only native subjects took part in the rating task. We used the same 30-second excerpts of all sales pitches as stimuli, thus controlling for verbal content. Between 19 (BRAZ) and 24 (GER) listeners participated per language, based on an online SoSci Survey setup. As we included 20 speakers per language, a total of about 400 ratings were collected per language. All listeners used headphones when rating the stimuli.

Perceived speaker charisma was rated based on the question “how much money would you in invest into the speaker’s idea on the scale from 0 to 1000 _?”. The _ symbol was replaced by each country’s individual currency symbol. We chose this task because previous studies proved that investment likelihood is closely associated with key attributes of charisma: decisive, performing, inspiring, and persuasive.

6.2. Results

Results are summarized in Figure 2. They show a clear and significant difference (p<0.001 according to an ANOVA) in investment between countries with a small gender gap (ESP, GER) on the one hand and countries with a big gender gap (UKR, BRAZ, CHIN, TURK) on the other. UKR, BRAZ, CHIN, TURK listeners would spend more money on male presenters, while investing on average less than half into female presenters (p<0.001). Interestingly, this applied even more so to female than to male listeners (esp. from UKR & TURK). By contrast, ESP and GER listeners would invest overall “only” 20-30% less into sales pitches held by women – and this applies equally to both female and male listeners (p>0.05). Therefore, the answer for our H2b is also positive.

7. Discussion

The present paper was about charismatic speech in different countries. Our aim was four-fold: 1) to examine acoustic correlates of charisma; 2) to search for an effect of speaker gender on the charisma ratings; 3) to check if a country’s gender (in)equality status shapes the production or/and perception of charismatic speech; 4) to relate the ancient, rhetorical history of persuasive speech to the modern concept of charisma so as to explain the difference in acoustic-prosodic features produced by the speakers in different countries. We found initial empirical evidence in line with 3 of our 4 hypotheses, i.e. H1 and H2a/b. Regarding the last hypothesis (H3), it is hard to give unambiguous answer. That is, compared to the speakers from the other 5 countries, Chinese speakers tended to use a narrower pitch range, a slower speech rate, less F0 variation and a breathier voice when given a charismatic speaking task. This accords well with the Chinese rhetoric tradition of oral restraint. However, representatives of Aristotle’s rhetoric like, for example, ESP speakers, created a similar results pattern. Also the representatives of Muslim rhetoric did not fully reflected the expected acoustic-prosodic charisma patterns, at least regarding the number of silent pauses.

8. Conclusions

All in all, UKR, TURK and BRAZ appear to use the same acoustic profile of charismatic speech. GER and ESP behaved similarly in rating speaker charisma through investment sums. CHIN charismatic speech is similar to that of ESP, but CHIN perceptual preferences strongly correlate with those of UKR, TURK and BRAZ listeners. The sample of 6 languages suggests that a country’s WEF/UN gender-gap score is correlated more strongly with speech-perception than with speech-production behavior. Additionally, a country’s gender gap seems to have a greater effect on charisma than a country’s rhetorical history.

Of course, both the production and the perception experiment rely on a limited set of prosodic parameters. Follow-up studies will extend that set, also aiming to determine their language/country/culture-specific relevance.

9. References


