What Makes a Speaker Sound Charismatic? Producing and Perceiving Charismatic Speech

Julia Hirschberg

November 2021

Columbia University

Collaborators

- *Andrew Rosenberg, Svetlana Stenchicova Stoyanchev, Wisam Dakka, Aron Wahl, Judd Sheinholtz, Fadi Biadsy, Rolf Carlson, Eva Strangert, Erica Cooper, Tomer Aharoni, Zixiaofan Yang, Nishi Cestero, Jessica Yin Huynh, Riku Tabata, Brandon Liang, Alina Ying…*


Detecting

What is Charisma?

- The ability to attract, and retain followers by virtue of personal characteristics -- not traditional or political office (Weber ‘47)
  - Political, religious or even business leaders: Gandhi, Hitler, Castro, Martin Luther King Jr., Steve Jobs….
  - ***Personalismo***
- What makes an individual charismatic? (Bird ‘93, Boss ‘76, Dowis ‘00, Marcus ‘67, Touati ‘93, Tuppen ‘74, Weber ‘47)
  - Their message? Their personality? Their speaking style?

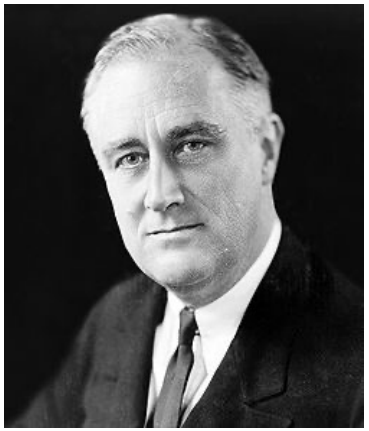
What makes Speech Charismatic?

- Circularly…
  - Speech that leads listeners to perceive the speaker as charismatic
- So…what aspects of speech and language contribute to perceptions of a charismatic speaker?
  - **Content** of the message?
  - **Lexico-syntactic features**?
  - **Acoustic-prosodic features**?
  - How do **facial expression** and **body gestures** contribute?

Why should we Study Charismatic Speech?

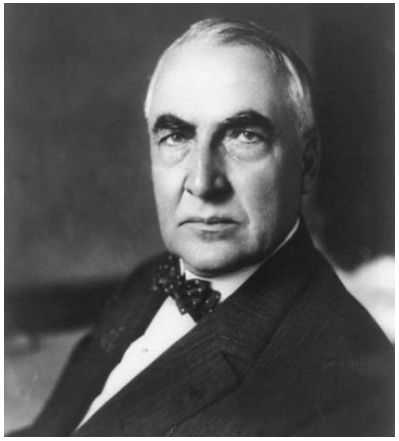
- Speech that leads listeners to perceive the speaker as charismatic is ***useful*** for multiple reasons:
  - Improving ***Text-to-Speech Synthesis***: charismatic speakers are much more appreciated by listeners, more trusted, more persuasive
  - ***Explaining success*** in political or other venues
  - Helping ***human speakers to improve*** their own speech production: to be more successful in startup pitches or just to produce better, more persuasive, more compelling talks…
  - Now for some historical examples...

1. Vladimir Lenin



2. Franklin D. Roosevelt

3. Mao Zedong



4. Warren G. Harding

5. Adolf Hitler

Our Goals

- To identify charismatic speakers
  - To understand which **aspects of spoken language** are correlated with listener-perceived charisma – and which ***other*** **personal characteristics**
  - To study charisma perception in **different modalities**
  - To compare charisma perception **across cultures**
  - To see how demographic and personality differences affect listeners' charisma perceptions
- And ultimately, to create more charismatic text-to-speech **voices** and to build an **interactive game** that can help speakers learn to be more charismatic

Early Experiments

- Collected tokens of charismatic and non-charismatic speech from a small set of politicians on a small set of topics
- Asked listeners to rate the 'The speaker is charismatic' and a number of other speaker traits (e.g. The speaker is …***boring, charming, persuasive,***...)
- Correlated listener ratings with **lexico-syntactic** and **acoustic-prosodic features** of the tokens to identify potential **cues to perceptions** of charisma

American English Perception Study

- Data: 45 2-30s speech segments, 5 each from 9 candidates for Democratic nomination for U.S. president in 2004

  - **Speakers**: Liberman, Kucinich, Clark, Gephardt, Dean, Moseley Braun, Sharpton, Kerry, Edwards
  - **Topics**: greeting, reasons for running, tax cuts, postwar Iraq, healthcare
  - **Genres**: stump speeches, debates, interviews, ads
  - 8 subjects **rated each segment** on a Likert scale (1-5) for 26 questions including whether they **agreed with the speaker** or not and whether they found the speaker's **message clear** (rating time avg. 1.5 h; min 45m, max ~3hrs)

How Much do Raters Agree with *Each Other*?

- Over **all statements**?
  - Using **Cohen's kappa** statistic with **quadratic weighting** to measure agreement betwen 0 and 1, mean $\kappa = 0.207$
- On **charisma**?
  - $\kappa = 0.232$ (8^{th} most agreed upon statement)
- By **token**?
  - No significant differences across all tokens although Edwards and Liberman tokens show significantly more agreement across all statements – why? Who is most **charismatic**? Least?
- By other speaker **traits**?
  - Individual traits demonstrate significantly different agreement (most: *The speaker is accusatory, angry, passionate, intense;* least: *The speaker is trustworthy, believable, reasonable, trustworthy*)

How do Raters Define Charisma?

- Which *other* traits are most closely correlated with the charisma? a *functional* definition

The speaker is enthusiastic	0.620
The speaker is persuasive	0.577
The speaker is charming	0.575
The speaker is passionate	0.543
The speaker is boring	-0.513
The speaker is convincing	0.499

Does Rater Agreement with a Speaker Correlate with Their Charisma Judgments?

- Testing earlier hypotheses (Weber 1947;Boss 1976; Dowis 2000): Whether a rater ***agrees with*** the speaker does correlate but not very highly with charisma judgments ($\kappa = 0.30$)

Does *Recognizing a Speaker* Correlate with Charisma Judgments?

- Raters were asked to identify which, if any, speakers they **recognized** at the end of the study.
- Mean number of speakers ***believed*** to have been recognized, 5.8
- Raters judged '**recognized**' **speakers** as significantly ***more charismatic*** than those they did ***not*** believe they had recognized *even when their recognition was not correct* (mean 3.39 vs. mean 3.30).

Does *Genre* or *Topic* Correlate with Judgments of Charisma?

- Recall that tokens were taken from **debates**, **interviews**, **stump speeches**, and **campaign ads**
  - **Genre *does*** influence **charisma** ratings (p=.0004)
    - Stump speeches were the most **charismatic** (3.38)
    - Interviews were the least (2.96)
- **Topic *does*** affect ratings of charisma significantly (p=.0517)
  - Healthcare > post-war Iraq > reasons for running > neutral > taxes

What makes Speech Charismatic: Lexical and Acoustic-Prosodic Features?

- ***Duration*** (secs, words, syls)?
- *More **personal**?*: Pronoun density
- *More **contentful**?*: Function/content word ratio
- ***Simpler***: Complexity: mean syllables/word (Dowis)
- *More or less **disfluencies**?*
- *More or fewer **repeated** words?*

- Min, max, mean, stdev ***F0*** (Boss, Tuppen)?
  - Raw and normalized by speaker
- Min, max, mean, stdev ***intensity**?*
- ***Speaking rate*** (syls/sec)
- ***Intonational*** features?:
  - Pitch accents
  - Phrasal tones
  - Contours

Findings: *Lexico-Syntactic* Correlates of Charisma

- ***Length***: Greater number of words ***positively*** correlates with charisma (r=.13; p=.002)
- ***Personal pronouns***:
  - Density of first person plural and third person singular pronouns (e.g. “we”, “he/she/it”) ***positively*** correlates with charisma (r=.16, p=0; r=.16, p=0)
  - Third person plural pronoun (e.g. “they”) density ***negatively*** correlates with charisma (r=-.19,p=0)
- ***Content***: Ratio of adjectives/all words ***negatively*** correlates with charisma (r=-.12,p=.008)
- ***Complexity***: Higher mean syllables/word ***positively*** correlates with charisma (p=.034)

- ***Disfluency***: greater % ***negatively*** correlates with charisma (r=-.18, p=0)
- ***Repetition***: Proportion of repeated words (a standard rhetorical device) ***positively*** correlates with charisma (r=.12, p=.004)

Findings: Acoustic-Prosodic Correlates of Charisma

- ***Pitch***: Higher F0 (mean, min, mean HiF0, over male speakers only) ***positively*** correlates with charisma (r=.24,p=0;r=.14 p=0;r=.20,p=0)
- ***Loudness***: Higher mean (normalized) rms and sdev of mean rms ***positively*** correlates with charisma (r=.21,p=0;r=.21,p=0)
- ***Speaking Rate***: Faster overall rate (voice/unvoiced frames) ***positively*** correlates with charisma (r=.16,p=0)
- ***Duration***: Longer duration correlates positively with charisma (r=.09,p=.037)

- Length of ***pause***: sdev ***negatively*** correlates with charisma (r=-.09,p=.004)
- **Pitch accent** types: !H* (downstepped contours) correlate positively with charisma while other types associated with question contours and uncertainty negatively correlate

So…What makes Speech Charismatic in American English?

- In Standard American English, speakers highly rated for charisma tend to be
  - Speakers highly rated for *enthusiasm, charm, persuasiveness, passionateness* and *convincingness* – and **not** *boring*
- Charismatic utterances tend to be
  - **Lexical**: *Longer*, with a *lower ratio of adjectives* to all words, *more first person plural and third person singular pronouns* and *fewer third person plurals*, *fewer disfluencies*, *more repeated words*, and *more complex words* than non-charismatic utterances

- **Acoustic-prosodic:** Charismatic utterances are *higher in pitch (mean, min)* with *more regularity in pause length, faster, louder,* and with *more variation in intensity*

Cross-Cultural Ratings of Charisma in American English and Palestinian Arabic

- Compare **ratings of Palestinian and English** corpora by ***American, Palestinian and Swedish listeners***:
  - Some acoustic-prosodic correlates are common across cultures, others are not -- but both native and non-native ratings reflect an apparent appreciation of both
  - ***Correlation between native and non-native rater of charisma is statistically significant for all language pairings on rating English and Arabic****, even though American and Swedish raters did not know Arabic and Palestinians did not know English – they shared acoustic-prosodic correlates*?

Rosenberg & Hirschberg,"Charisma Perception from Text and Speech," *Speech Communication* 2008.

Later Studies

- **Politicians**:
  - 1 Italian (Signorello et al., 2012); another Italian and 1 French (Derrico et al., 2013); 1 Irish (Cullen et al., 2014)
- **Business leaders**:
  - 143 business executives (Weninger et al., 2012)
  - Steve Jobs and Mark Zuckerberg (Mixdorff et al., 2018) (Niebuhr and Gonzalez, 2019): Jobs: more **charismatic**: higher F0 level, larger F0 range, more variability in speech, clearer pronunciation
- **University lecturers**:
  - 2 male native speakers of English reading same text which is then manipulated for f0 and speech rate and rated by 16 other native speakers, finding that low F0 range and low speaking rate were ranked as most **charismatic** (Berger et al., 2017)

- **Female charismatic speech**
  - Business leaders: 2 females (Oprah Winfrey, Ginni Rometty) produced **stronger** acoustic **charisma** cues but were still judged to be similar in **charisma** to the single male speaker (Steve Jobs) (Novák-Tót et al., 2017)
  - Charisma training system: Female speakers start with significantly lower **prosodic-charisma scores** in short "pitches" than male speakers, judged by Pascal, an automatic training and scoring system that analyzed acoustic-prosodic cues on 21 males and 16 females taking the 4h version and 20 males and 15 females in the 12wk version (Niebuhr et al., 2019)
  - **Charisma cues** and scoring **metrics** in both taken from previous literature on male speech only

But Little Research on Broader Demographics of Charisma Perception and Production

- Do *raters'* **demographics** or **personality** information affect their **charisma** perception?
  - *Gender, level of education, sexual orientation, gender attraction, personality* (TIPI) scores
- What other attributes are associated with **charisma** by different groups?
- How do **raters'** ***own productions*** of **charismatic** vs. "normal" speech compare with speech they rate as more **charismatic** in acoustic-prosodic features?

"What Makes a Speaker Charismatic? Producing and Perceiving Charismatic Speech," Zixiaofan Yang, Jessica Huynh, Riku Tabata, Nishmar Cestero, Tomer Aharoni, Julia Hirschberg. Speech Prosody, May 2020.

New Data Collection

- First large **gender-balanced** study of **charismatic** speech of non-celebrities, non-politicians
  - 60 20sec clips from online talks of 30M/30F spkrs

  chosen as likely rated **charismatic**, **not charismatic**, or somewhere in between), each rated by 15-20 raters
- **Amazon Mechanical Turk** ratings:
  - 97 English-speaking raters, surveyed for gender, sexual preference, level of education, TIPI personality traits to determine whether demographic information influences charisma opinions?
  - Also asked to provide 2 voice samples of the same text: one normal, one "***charismatic***" – how does the rater's speech influence their perception of others?

Ratings on 18 Traits

Speaker Traits	1 Completely Disagree	2 Somewhat disagree	3 Neither Agree/disagree	4 Somewhat Agree	5 Completely Agree
Confident	○	○	○	○	○
Sincere	○	○	○	○	○
Extroverted	○	○	○	○	○
Ordinary	○	○	○	○	○
Weak	○	○	○	○	○
Fluent	○	○	○	○	○
Intelligent	○	○	○	○	○
Uncertain	○	○	○	○	○
Introverted	○	○	○	○	○
Persuasive	○	○	○	○	○
Enthusiastic	○	○	○	○	○
Reasonable	○	○	○	○	○
Cold	○	○	○	○	○
Eloquent	○	○	○	○	○
Boring	○	○	○	○	○
Trustworthy	○	○	○	○	○
Charismatic	○	○	○	○	○
Lively	○	○	○	○	○

Some Examples for You to Rate for *Charisma* (1-5)

- One  2.31
- Two  4.50
- Three  4.10
- Four  2.06

Raters' Demographics

- **Gender**:
  - 60 female, 36 male, 1 preferred not to say; 68 heterosexual, 11 bisexual, 16 homosexual; 42 attracted to females; 65 attracted to males
- **Education**:
  - some school (1), high school (21), associates (19), BA/BS (45), MA (10), PhD (1)
- Average scores on the **TIPI Big-Five personality** dimensions (1~7):
  - 5.12 openness, 5.54 conscientiousness, 3.70 extroversion, 5.39 agreeableness, 4.91 emotional stability -- distribution skewed toward higher scores for all dimensions except extroversion

Questions to Answer

- How do raters define **charisma** in terms of other speaker ***traits***?
- Does ***genre*** of recording influence **charisma** ratings?
- Does ***speaker gender*** influence raters' **charisma** ratings or their ratings on other traits?
- What are the ***acoustic-prosodic*** and ***lexical*** properties of speech rated as **charismatic**?
- Does a rater's ***demographic information*** and ***personality*** characteristics influence their ratings?
- Does ***raters' own speech*** correlate with their **charisma** ratings of others or their own demographics/ personality?

Analysis and Results

- How do raters define **charisma** in terms of other speaker **traits**?
- Correlations between **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

- Consistent across both **speaker** and **rater genders**

- Inter-rater **agreement** on the traits
  - **Charisma**: 0.296
  - 5 most and 5 least agreed-upon 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

- **Higher activation** traits are more agreed upon, and **lower activation** traits are more open to interpretation

Influence of Genre on Ratings

- Does the **genre** of the recording influence **charisma** ratings?
  - 14 **interviews**, 19 educational **lectures**, 27 **talks** to more general audiences
  - **Interviews** are less charismatic than educational **lectures** ($p = 0.03$) and **talks** ($p = 0.001$)
  - Why? The **goals** of the speech are different?

Influence of Gender on Ratings

- ***Female*** speakers achieved a higher average **charisma** score than male speakers, although not significantly different ($p = 0.153$)
- ***Male speakers*** were rated as less sincere, less fluent, and less extroverted, but more boring and more introverted than females
- The lower **charisma** score of males may be due to **genre** (more **lectures** for males) and not gender…

Acoustic-Prosodic Properties of Charismatic Speech

- **Normalized** male/female speech features using the mean values for American English speakers
- **Gender-specific *positive* correlates**:
  - **Female speakers**: mean intensity, standard deviation in pitch
  - **Male speakers**: mean pitch, speaking rate, standard deviation in pitch
- Positive correlates for **both genders** with **charisma**: Mean intensity, mean pitch, speaking rate, standard deviation of pitch

– **Charismatic** speech is louder, higher, faster, and with greater fluctuation in pitch ***as we found in the earlier studies***

Lexical Properties of Charismatic Speech

- ***Linguistic Inquiry and Word Count*** (LIWC)
  - Positively correlated over both genders: **interrogative words**
  - Negatively correlated over both: **first-person pronouns, negative emotion words, sadness words, discrepancies, words of feeling**
  - Speakers that use **negative emotional words** were rated as less **charismatic** regardless of gender

Ratings and Rater Demographics

- A few interesting **findings**:
  - **Male raters** rated speakers as weaker and colder than females did
  - Raters with **higher education levels** rated speakers as more eloquent and lively but also less intelligent
  - Raters with **higher scores in *openness, emotional stability, conscientiousness, and agreeableness*** tended to rate speakers higher in **charisma** and traits positively correlated with **charisma** and lower in traits that negatively correlated with **charisma** – ***using themselves as a reference?***

Raters Own Charismatic and Non-charismatic Speech

- Comparing raters' own ***normal*** speech with their production of ***charismatic*** speech (paired t-tests)
  - Raters significantly increased their **mean intensity**, **mean pitch** and **standard deviation of pitch**, and decreased their **HNR** when asked to be **charismatic**
  - Quite **similar to how raters rated our speakers' voice clips** – but **male raters** changed their voices more than **females** did

Most Recent Research

- ***Politicians' speech***: 24 of 25 Democratic candidates running for president in 2020
  - 4 genres: ***Debates, Campaign Ads, Interviews, Stump Speeches***
  - Collected ratings of 294 speech clips from AMT using processes similar to those in our cross-gender and early politician research

A Few Interesting Results…

- AMT ratings were collected just before and after election day which clearly influenced opinions
  - Raters were **divided by political leaning**: 28 liberals, 13 conservatives, and 15 moderates and these differences did correlate with a number of their ratings
  - **Charisma** ratings were again **positively correlated with similar acoustic-prosodic features across all gro**ups, but there was very little agreement otherwise
  - **Significant difference in how female vs. male speakers were viewed**, supporting prior findings that words indicating strength and toughness are desirable in male candidates but much less so in females
  - However, in general, we concluded that politicians' **charisma** is currently a very divisive issue in the U. S. today ☹

But Sounding Charismatic can still be important for some….

- An **example from the past**….listen to the first voice and compare it to the second…

Why is Charisma so Important?

- These are from the same person….

How Useful is Charismatic Speech?

NYTimes, 11/23/21: "Using charisma, poise and a smattering of scientific jargon, Elizabeth Holmes persuaded investors to give her nearly $1 billion to build Theranos, her blood testing start-up. That all came crashing down in 2018, after the company's technology and business dealings were revealed to have major problems.

"On Tuesday, Ms. Holmes used those same techniques to try to convince a jury that she was not guilty of fraud."

- ***We shall see…***

Other Current Projects

- Identifying false information on COVID-19 and climate change in Twitter and the motivations behind it
- Discovering radicalizing aspects of far right and far left group videos
- Detecting hate speech directed at women journalists
- Deceptive and trusted speech with LieCatch training
- Producing appropriate acoustic-prosodic features in TTS
- Identifying emotion in speech across languages
- Detecting and producing empathetic speech
- Correcting alignment errors in the SWBD Corpus

Thank you!