

Humble Voices in Political Communication: A Speech Analysis Across Two Cultures

Francesca D'Errico[1(✉)], Oliver Niebuhr[2], and Isabella Poggi[1]

[1] Fil.Co.Spe Department, Roma Tre University, Rome, Italy
{francesca.derrico,isabella.poggi}@uniroma3.it
[2] Centre for Industrial Electronics, University of Southern Denmark, Sonderborg, Denmark
olni@sdu.dk

Abstract. Classical works on affective and persuasive computing have pointed out that technologies whose goal is to influence the user's need to communicate in a charismatic or dominant style, neglecting the potential persuasive strength of other styles, such as a humble style. Humility can be defined as an epistemic and interactional stance aimed at communicating the person's attitude towards an object, topic, or the interlocutor. Previous studies on persuasion in politics conducted multimodal analyses to identify those postures, prosodic features, gaze patterns, and facial expressions that convey humility. In this study, we describe the features of humility associated with a speaker's tone-of-voice, by comparing two pairs of popular politicians from USA (Obama vs Trump) and Italy (Gentiloni vs Salvini) during TV interviews. Results on the acoustic side point out clear differences in the speech characteristics of humble vs. dominant politicians: Gentiloni's and Obama's speech is characterized by short utterances, more hesitations and disfluencies and fewer stressed words. The two speakers' loudness levels are lower, and their voice qualities either more sonorant or even breathier. We discuss our results with respect to their implications for persuasive processes which can be associated not only to the speakers' features but also to the speaker-audience personality-trait matching.

Keywords: Humility · Multimodal communication · Voice · Speech analysis

1 Introduction

Within studies on persuasion, the notion of humility has been neglected since influence over others has been generally viewed as a process where a 'dominant' or a charismatic leader generally uses different maneuverings to show he has more power than his opponent (Fogg 2002; Poggi and D'Errico 2010; D'Errico et al. 2013; Niebuhr and Ochoa 2019). From this point of view, persuasion research has neglected the potential individual differences related to the preference of a dominant vs. humble speaker. In recent studies, in fact, it has emerged that individuals with a low orientation to social dominance, high self-esteem and high level of moral relevance experience much more positive emotions (i.e. joy, enthusiasm) when they listen to communication phrased in a humble rather than a dominant way (D'Errico, forth.). Therefore, in the field of

© Springer Nature Switzerland AG 2019
S. Misra et al. (Eds.): ICCSA 2019, LNCS 11620, pp. 361–374, 2019.
https://doi.org/10.1007/978-3-030-24296-1_30

persuasive technologies (Franzoni et al. 2019) on bodily and linguistic signals that promote a persuasive process in humble communication becomes more and more relevant. In this regard, a definition of humble stance will first be presented by referring to a socio-cognitive model of persuasion in terms of goals and beliefs (Poggi 2007). Subsequently, recent studies on the multimodality of humble communication (D'Errico and Poggi 2017; 2019) will be reviewed, highlighting how the body communicates humility. The last section will present an empirical study focused on the humble speech features as compared with the dominant ones, through advanced speech detection techniques. The speech analysis will also describe potential cultural differences since the audios analysed are selected from two Italian and two American politicians. In particular we present an acoustic analysis by means of the speech-signal processing software PRAAT (Boersma 2001), extracting eleven acoustic prosodic parameters concerning:

(1) Pausing and speaking, measured in the time domain of the signal;
(2) Speech melody, measured in terms of f0 in the spectrum of the acoustic signal and
(3) Tempo and voice.

Furthermore, a second analysis will describe conversation dynamics including four additional parameters like number of disfluencies, overlapping turns, mean duration of overlapping and number of competitive turn taking. Results will be discussed in view of understanding the differences between dominant and humble leaders' voice and its relation to the leaders' charisma.

2 The Humble Stance

In psychological terms, humility is a kind of interactional stance, namely, a way of interacting with others while admitting one's own limitations, showing openness to new ideas and means, giving others a voice, and recognizing their merits (Weidman et al. 2018).

On the opposite side, studies on political persuasion have generally pointed out how a charismatic and trustworthy politician needs to be acknowledged as moral, competent but also dominant. In addition to morality and competence, in fact, the feature of dominance seems to be effective in persuading an audience (Poggi and D'Errico, 2010). Dominance has been defined as a relational construct implying a power comparison (to have "more power than" another); but also the communication of power—the fact of producing signals conveying one's dominance—is very important, because communicating power is a way to maintain it or even to acquire more of it (Castelfranchi 2003). Poggi and D'Errico (2010), in their multimodal analysis of dominance strategies, have shown that politicians sometimes use 'blatant' dominance strategies, clearly characterized by aggressiveness or power display, like using speech acts of criticism or accusation, or speaking aloud; but they can also exploit 'subtle' dominance strategies, that at first sight do not show a clear exhibition of force or power, but are nonetheless more indirect ways to show dominance, like showing a calm strength or ignoring the other.

The preference for the dominant communication in its explicit forms (e.g., discrediting the opponent) is higher when people show adherence to power, security and social dominance orientation (D'Errico and Poggi 2012a; b).

Dominance in some cases plays a role also in political charisma, that set of features of a politician which allow him to influence followers' actions and emotions not by the use of force but through admiration and identification. As shown in recent studies (Signorello et al. 2012; D'Errico et al. 2013), from a perceptual point of view the persuasive strength of charismatic features is correlated with features of dominance.

On the contrary, a humble stance can be defined as a multimodal public act, performed interactively through verbal and bodily communication, by which a person positions oneself with respect to the object of a communicative interaction and also with respect to one's interlocutor. In this sense, a humble stance of a politician can be expressed either by communicating his 'horizontal', 'equal' position toward his interlocutor or by acknowledging one's uncertainty (Vincze et al. 2012), limits and flaws, and thus the consequent possibility of making errors (limit awareness).

According to the qualitative analysis by D'Errico and Poggi (2017) aimed at finding out the features of Humility, the humble stance is a "realistic" approach, an ability to keep one's feet firmly on the ground, in the awareness of being always fallacious and never perfect. Limits awareness implies a feeling/communication of equality to others (Equality feature), not to feel superior to them, not to display any superiority, and if having some power over others, not to take advantage of it (Non-Superiority feature). This stance has behavioral consequences in that the humble person shows being on the same level as others, as well as treating them as equal to oneself (Familiarity), and being empathic to them (Empathy), which implies Care and Attention to the other; hence giving an impression of Altruism, of being oriented to other people more than to objects or to oneself.

The basic feature of Realism entails, from an expressive point of view, not crediting any relevance to any external tinsel or symbolic ornament like status symbols or status tout court (Essentiality), which results in features of Informality and Sincerity, not caring anything but the real substantive value of people.

Such definition of humility from a cognitive perspective shows how the main goal of a humble person is to be "like others", not more and not less, thus for pursuing this goal she does not show nor make appeal to his/her own power, superiority in terms of status, knowledge, merits, contributions, virtues and capabilities. From a communicative point of view s/he considers important "not to put her/himself first" but rather attributes a positive value to a larger dimension of belonging (e.g., others, group, organizations, party) and focuses on the problem rather than on the person who did something. Such horizontal perspective of the humble person focused on others and the group, leads to emphasizing elements of similarity, familiarity, and informality.

3 Charisma and Its Multimodal Communication

In previous works, studies in the psychology of persuasion (D'Errico et al. 2013) and acoustic analysis of voice (Signorello et al. 2012) have investigated the notion of charisma while trying to find out the features of charismatic voice. Charisma, first

defined by Weber (1920) as an "extraordinary quality" of a person who is believed to be endowed with superhuman properties, therefore inducing people to acknowledge him as a leader and to develop a cult of him, can be defined, in terms of a socio-cognitive model, as a set of internal features of a person that, when manifested by some external displays – physical traits or behaviors displayed in various modalities – trigger a set of emotions in other people, that induce them to pursue some goals not through coercion but voluntarily and willingly, while feeling involvement, passion, and enthusiasm in what they do together or on behalf of that person. These internal features are displayed in the person's multimodal behavior: his words, gestures, gaze, facial expression, posture, and voice. Inspiring to the cognitive model of persuasion (Poggi 2007), according to D'Errico et al. (2013) the internal features include personal characteristics that endow the person with crucial dimensions of persuasiveness: *emotional intelligence* (the skill of feeling emotions himself, of displaying them to others, and to be empathic with others' emotions) and *emotional induction* (the skill to transmit emotions to others through expressivity and contagion), *benevolence* (an attitude of sociability and inclusiveness: being people-oriented, making followers feel "similar" to him and "together" with him); *competence* (both physical and intellectual capacities: vision, creativity, foresight, and communicative skills), and finally *dominance* (capacity of winning in competitions, not to submit but to dominate).

The external features that display such charismatic internal features of charismatic leaders were analyzed in the words, gestures and postures of Benito Mussolini (Poggi 2017), in the voice of other Italian and French politicians (D'Errico et al. 2013; Signorello et al. 2012), where the parameters of pitch, jitter and shimmer were taken into account, and in the voice of leaders in organizational settings (Niehbur et al. 2016) showing that Steve Jobs's voice, compared to other reference samples, stands out against them in almost all key features of charisma, including melody, loudness, tempo and fluency, thus producing significant quantitative differences in influence potential when addressing customers and investors in his speeches.

While analyzing the change in acoustic parameters of the voice of Umberto Bossi due to a stroke (Signorello et al. 2012), it was found that his charisma did not totally fade away from before to after the stroke, but it changed in type. Three types of charisma were singled out, stemming from different combinations of the charismatic dimensions: an Authoritarian – Threatening one, with prevailing features of dominance and aggressiveness; a Competent-Benevolent one, with the prevalence of skills display and empathy towards followers; and a Proactive-Attractive one, characterized by optimism, orientation to future, and a seductive attitude. Charismatic leaders differ from each other due to their prevailing type of charisma: Mussolini mainly exploits an Authoritarian-Threatening one, with shifts to Proactive-Attractive (Poggi 2017); Bossi shifts from an authoritarian-threatening one before the stroke to a more competent-benevolent in his voice after the stroke. Another study (D'Errico et al. 2013) analyzing the effects of higher or lower pitch and longer or shorter pauses of leaders on the followers found out a cultural difference between Italians and French in the perception of charisma and the interpretation of its signals: both Italians and French appreciate a Proactive-Attractive leader more than a Calm-Benevolent leader, though with slight differences. Both perceive a Calm-Benevolent charisma while listening to short pauses, but the perception differs for pitch. This might mean that they differ in the interpretation

of a factor of "human sociability" (a Calm-Benevolent charisma). Italians attribute it more to a high-pitch voice with shorter pauses, whereas the French prefer low-pitch voice and long pauses.

4 The Multimodal Communication of Humble Leadership

D'Errico and Poggi (2017) analysed the multimodal communication of humble leaders. Here the point is to find out the external features in the leader's words, gestures, posture, facial expression that are a cue to the internal features outlined above: equality, non-superiority, empathy, and so on. To this aim, fragments of communication of humble leaders were analysed in terms of an annotation scheme in which, for each body signal (gesture, posture, gaze item) its meaning is annotated, and its underlying humility feature. In the analysis (D'Errico and Poggi 2017), the President of Uruguay José Mujica is described as dressed with a *simple shirt* – a signal of INFORMALITY – *leaning forward* as if communicating a desire to get closer, to create FAMILIARITY with the interlocutor; his *head* is *lowered* and directed *straight to the camera* but, like gaze

Table 1. Bodily features of humility

Parameter	Signal	Meaning
Head	Bent	Humble
	Leaning forward	Request Or other-oriented
	straight	Authoritative
	Tilted backward	Proud/arrogant
	Head canting	Well-disposed/listening Or Ironic/skeptical
Chin	Downward on chest	Humble, sad, or ashamed
	Downward forward	Request Or other-oriented (Well-disposed towards the other)
	Default	Authoritative
	Upward	Proud/arrogant
Trunk	Forward relaxed	Humble or modest
	Leaning forward tense	Other oriented/humble
	Erect	Authoritative
	Strutting (straight tense)	Arrogant
	Slouched (backward relaxed)	Careless Or arrogant
Shoulders	Closed forward	Humble or modest
	Bottle shoulders	Sad Or humble
	Horizontal	Self-confident/Authoritative
	Backward	Arrogant Or distant
Gesture amplitude	High	Aggressive (possibly mocking exaggeration
	Default	Authoritative
	Within shoulder width	Humble

direction, from down upward, *both positioning him in a down-up position, conveys the meaning that he does not judge or command over the potential listener (NON- SUPERIORITY). His gaze, by* raised and frowning eyebrows*, expresses worry, hence concern about the interlocutor's future (EMPATHY), while* nose wrinkles tightened lids *and* tightened lips*, that convey anger, communicate the importance of the subject by raising attention.*

Based on such multimodal analysis on several fragments of humble leaders' communication, D'Errico and Poggi (2019) draw a tentative list of correspondences between head, hands, and trunk movements and positions with the features of humility or, to the contrary, with opposite features of superiority or distance (Table 1).

5 The Voice of Humble Politicians Across Two Cultures

The main aims of our work are: 1. to find out the characterizing features of humility in the voice of humble politicians; 2. to test if they are significantly different from features of dominance in dominant-aggressive politicians; 3. to assess possible differences in the external features of humility in two different cultures, American and Italian.

A final aim of our study is to try to disentangle the complex relationship that holds among dominance, humility, and charisma. Is charisma perceived only in dominant leaders, or is there a charismatic humble leadership?

5.1 Comparing American and Italian Dominant and Humble Leaders

To search real cases of political communication by dominant and humble leaders, in a preliminary study we asked participants to list two humble and two dominant American and Italian politicians. Then we chose four politicians (two humble and two dominant) among those most frequently pointed by our participants. In order to explore potential cultural differences we chose two politicians from Italy, Paolo Gentiloni (Prime Minister of the Italian Republic for two years, since 2016 to 2018: the humble one) and Matteo Salvini (Minister in 2018 and 2019; the dominant one), and two from Usa, Barack Obama and Donald Trump, humble and dominant, respectively.

5.2 Method: Speech Material and Analysis

Speech samples of the four politicians were extracted from public YouTube videos. Care was taken that all samples were comparable: all the politicians were interviewed in a well established TV talk show by a male interviewer. Both politician and TV host sit on a chair face-to-face about 2 m away from each other. About 5 min of speech were extracted per politician from the middle of each interview, at a point which was largely free from audience background noise and applause. All turns (i.e. coherent speech contributions like questions or comments) of the talk-show host were removed from the extracted speech sample so that only the respective politician's speech could be analyzed.

The first analysis was conducted on an acoustic basis, using three existing scripts that were written for the PRAAT speech-signal processing software (Boersma 2001) by

de Looze and Hirst (2008), de Jong and Wempe (2009), and Xu (2013). Eleven acoustic-prosodic parameters were included in the analysis, covering different characteristics of the tone-of-voice of each politician:

- Pausing and speaking, measured in the time domain of the signal

  (1) number of silent pauses (per minute, min)
  (2) mean duration of silent pauses (milliseconds, ms)
  (3) mean duration of inter-pausal speech units (ms), i.e. everything in between two audible pauses with durations >200 ms, see Walker and Trimboli (1982) and Heldner (2011)

- Speech melody, measured in terms of the fundamental-frequency or f0 in the spectrum of the acoustic signal

  (4) mean pitch level (f0 mean, in Hz),
  (5) mean lower limit of the speaker's pitch range (f0 minimum, in Hz)
  (6) mean upper limit of the speaker's pitch range (f0 maximum, in Hz)
  (7) mean pitch range (f0 min-f0max in number of octaves, semitones, st),
  (8) pitch variability (f0 standard deviation, in Hz),

- Tempo and voice

  (9) net speaking rate (in syllables per second, pauses excluded)
  (10) voice quality in terms of the Hammarberg index (in dB)
  (11) mean dispersion of formants 1–5 (in Hz)

The Hammarberg index is the acoustic-energy difference between two ranges of the speaker's voice spectrum (0–2 kHz and 2–5 kHz). Thus, the index determines how much energy a speaker's voice loses towards higher frequencies (Hammarberg et al. 1980). Louder and more powerful voices have a higher Hammarberg index and softer and thinner voices a lower one. Moreover, the index was found to vary systematically with the perceived emotions and traits of a speaker, with smaller values characterizing angry or urging speakers and/or speakers who sound more thrusting and powerful (Liu and Xu 2014; Nordström et al. 2017). Similarly, the mean dispersion of the first 5 formants is an acoustic correlate of the perceived body size of the speaker. The larger the volume, size, and weight of the speaker's body (i.e. the more imposing and awing the body), the smaller the value of the formant dispersion (Fitch 1997). Xu et al. (2013) found, accordingly, that formant-dispersion values are also related to emotional speech. Angry speakers have a higher formant dispersion than, e.g., happy speakers.

All measurements were taken per inter-pausal speech unit, with PRAAT's default analysis settings. Mean values were formed across all inter-pausal units of a speaker. The total number of inter-pausal units was 78 for Trump, 96 for Obama, 66 for Salvini, and 74 for Gentiloni. Thus, 314 measurements were taken per parameter. The measurements taken per speaker varied between 990 (Salvini) and 1,440 (Obama). The entire acoustic analysis comprised 4,710 individual measurements.

The second analysis concerned the conversation dynamics within the political interviews. It was conducted on an auditory basis by a trained phonetician (2nd author), partly because autonomous computer-based scripts were unable to perform the

requested measurements, and partly because the acoustics of the TV recordings made precise acoustic measurements difficult. Note that the auditory analysis was done on the basis of the entire TV interview. It included the following four additional parameters: (12) number of disfluencies, including filled pauses (tokens per min), (13) Number of overlapping turns of politician and TV host (tokens per min), (14) Mean duration of turn overlap (in seconds, s), (15) Number of competitive turn taking events on the part of the politician (tokens per minute). That is, the politician interrupts the moderator and takes his turn or continues talking while the moderator tries to interrupt, cf. Weilhammer and Rabold (2003) for a typology of turn-taking events.

Statistical Analysis
All parameters except for the relative frequencies (i.e. number of tokens per time unit, applies to parameters 1, 12, 13, 15) were statistically analyzed in separate two-way ANOVAs based on the two fixed factors STANCE (humble vs dominant) and COUNTRY (USA vs Italy). As there were only two levels per factor, additional multiple comparisons tests among factor levels were omitted. The relative frequencies were analyzed in terms of individual Chi-squared tests based on a 2×2 contingency tables per parameter.

5.3 Results

Table 2 summarizes the results of the acoustic and auditory analyses of the four politicians. The presentation of the results will start below with the factor STANCE (Sect. 5.3.1), i.e. the differences in interview behavior and tone-of-voice between humble and dominant politicians. The subsequent Sect. 5.3.2 briefly addresses the differences related to the factor COUNTRY. In Sect. 5.3.1, key results of Table 2 are additionally illustrated by figures.

5.3.1 STANCE: Humble vs Dominant Politicians

First, note that all significant differences between humble and dominant politicians were consistent across the two countries, i.e. they apply in the same way to both Italian and American speakers. In terms of the speaking and pausing behavior, this means that the humble politicians Obama and Gentiloni interspersed their turns in the TV interview with more pauses (χ^2 [1] = 26.3, $p < 0.001$) and longer pauses ($F[1,310] = 15.4$, $p < 0.001$) than Trump and Salvini. Trump and Salvini, on the other hand, talked a lot longer before making a (shorter) pause ($F[1,310] = 44.6$, $p < 0.001$), see Fig. 1.

With respect to speech melody, the results show that, compared to the humble politicians, the dominant politicians' speeches were characterized by a higher pitch level ($F[1,310] = 7.4$, $p = 0.007$). The same applies to the pitch minimum ($F[1,310] = 5.6$, $p = 0.02$) and pitch maximum ($F[1,310] = 10.9$, $p = 0.001$). In combination, these pitch differences point to a change of the entire pitch register, which seems to be generally shifted upward for dominant as compared to humble politicians, see also Fig. 2. The pitch range stays the same within the higher register of dominant politicians and, thus, does not differ significantly from that of the humble politicians. However, the pitch variability differs and is higher for the dominant politicians ($F[1,310] = 32.5$, $p < 0.001$).

Table 2. Results obtained per politician for the 15 acoustic and auditory parameters, organized in 2 × 2 fields representing the factors STANCE (horizontal) and COUNTRY (vertical); IPU = Inter-pausal Unit; N = 78 for Trump, N = 96 for Obama, N = 66 for Salvini, N = 74 for Gentiloni. Light grey 2 × 2 fields indicate significant differences only for the factor STANCE. Dark grey 2 × 2 fields indicate significant differences of both factors STANCE and COUNTRY.

Parameter types	Measures	Politicians (dominant)	Politicians (humble)	Measures	Politicians (dominant)	Politicians (humble)
Type 1 parameters: Speaking &Pausing	Silent pauses/min	Trump 14.7	Obama 21.4	ø silent pause dur. (ms)	Trump 559	Obama 670
		Salvini 10.3	Gentiloni 16.4		Salvini 406	Gentiloni 700
	ø IPU duration (ms)	Trump 3,424	Obama 2,068			
		Salvini 5,250	Gentiloni 2,923			
Type 2 parameters: Speech melody (f0)	ø Pitch level (Hz)	Trump 132	Obama 93	ø Pitch range (octaves)	Trump 1.9	Obama 1.9
		Salvini 140	Gentiloni 122		Salvini 1.7	Gentiloni 2.0
	ø Pitch min (Hz)	Trump 71	Obama 55	ø Pitch variability (sd, Hz)	Trump 30.9	Obama 20.5
		Salvini 81	Gentiloni 70		Salvini 39.7	Gentiloni 23.0
	ø Pitch max (Hz)	Trump 270	Obama 208			
		Salvini 330	Gentiloni 247			
Type 3 parameters: Tempo & Voice	ø Ham.brg. index	Trump 23.2	Obama 16.2	ø Formant dispersion (Hz)	Trump 594	Obama 870
		Salvini 17.6	Gentiloni 14.9		Salvini 768	Gentiloni 852
	ø Net speaking rate (syll/s)	Trump 5.6	Obama 4.4			
		Salvini 5.5	Gentiloni 4.8			
Type 4 parameters: convers. dynamics	Disfl., filled pauses/min	Trump 9.4	Obama 3.6	Turn-overl duration (s)	Trump 2.1	Obama 0.4
		Salvini 7.9	Gentiloni 5.5		Salvini 4.5	Gentiloni 1.7
	Overlapping turns/min	Trump 2.8	Obama 0.4	Comp. turn taking/min	Trump 1.8	Obama 0.2
		Salvini 4.2	Gentiloni 2.2		Salvini 2.6	Gentiloni 0.9

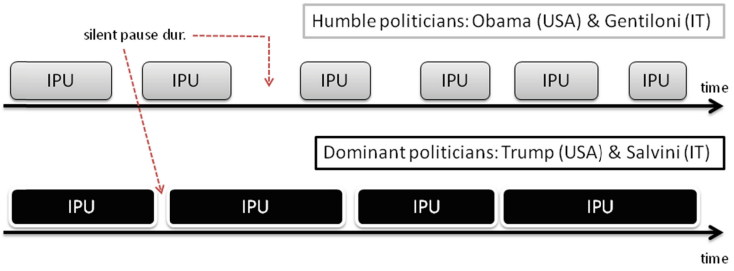

Fig. 1. Illustration of Type-1 parameter results: pausing & speaking

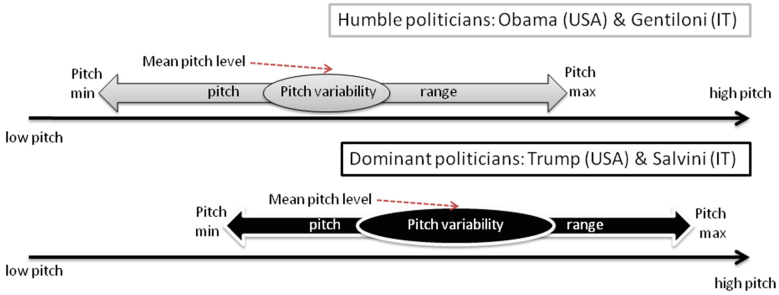

Fig. 2. Illustration of Type-2 parameter results: Speech melody

Regarding the analyzed parameters of tempo and voice, the dominant politicians spoke faster than the humble politicians ($F[1,310] = 4.6$, $p = 0.03$). We also found the Hammarberg index to be higher for the dominant than for the humble politicians ($F[1,310] = 11.1$, $p = 0.001$). The humble politicians' speech, on the other hand, was characterized by higher formant-dispersion values ($F[1,310] = 23.4$, $p < 0.001$).

Finally, the auditory analysis of conversation dynamics within the interviews adds two aspects to the overall results pattern. First, the dominant politicians Trump and Salvini produced significantly more disfluencies and/or used more filled pauses in their speech than the humble politicians Obama and Gentiloni (χ^2 [1] $= 5.2$, $p = 0.02$). Second, compared to the humble politicians, the data of their dominant counterparts reflect the strong will to control the conversation, i.e. to take the turn, to keep the turn, and to continue talking when the TV host tries to change the topic or go deeper into a certain argument, see also Fig. 3. Accordingly, overlapping turns (χ^2 [1] $= 4.1$, $p = 0.04$) and competitive turn taking (χ^2 [1] $= 19.6$, $p < 0.001$) are much more frequent for dominant than for humble politicians, and the duration of overlapping turns is also longer for dominant politicians ($F[1,310] = 56.7$, $p < 0.001$). Actually, also the fewer and shorter pauses of dominant politicians, and their preference for repaired over carefully pre-planned speech (larger number of disfluencies) and for filled

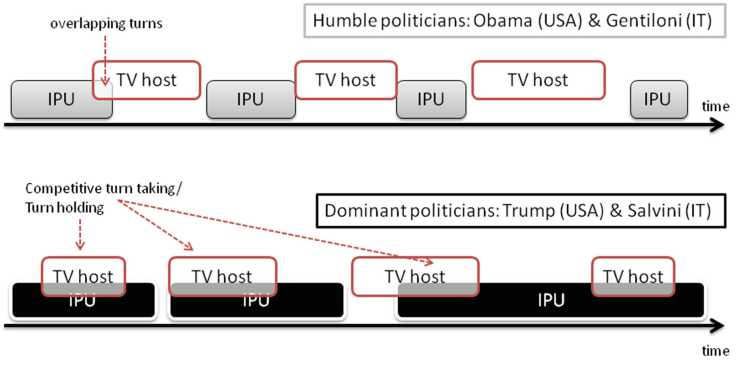

Fig. 3. Illustration of Type-4 parameter results: Conversation dynamics.

over silent pauses can be interpreted as reflecting the will to control the conversation by maximizing talking time and avoiding, by all means, silent intervals that could invite the TV host to start speaking.

5.3.2 COUNTRY: Italy vs USA

Table 2 shows that 66% of the analyzed parameters (i.e. 10 out of 15) also differ significantly as a function of COUNTRY. However, as was stated in Sect. 5.3.1, these differences are only quantitative and not qualitative in nature. That is, Trump (dominant) and Obama (humble) differ from each other in the same way as Salvini (dominant) and Gentiloni (humble), but on either higher or lower parameter levels and, for some parameters, also to a different degree. The latter is the reason for significant interactions STANCE*COUNTRY that emerged (all with $p < 0.001$) in the ANOVAs of three parameters: inter-pausal-unit duration, Hammarberg index, and turn-overlap duration. However, since all interactions are ordinal, we disregard them in the presentation of the results and focus on the main effects only.

For example, besides the differences due to STANCE, the two Italian politicians generally pause less often in their speech (χ^2 [1] = 14.3, $p < 0.001$) and speak in longer inter-pausal units than the two American politicians ($F[1,310] = 19.0$, $p < 0.001$). The whole speech melody (i.e. pitch register) is also produced at a higher level by the Italian than by the American politicians (pitch level: $F[1,310] = 12.8$, $p < 0.001$; pitch minimum: $F[1,310] = 3.9$, $p = 0.04$; pitch maximum: $F[1,310] = 9.3$, $p = 0.001$). The American politicians have a higher Hammarberg index, though ($F[1,310] = 8.8$, $p = 0.002$). Finally, our data suggest that political TV interviews in Italy are generally characterized by a higher conversation dynamics than TV interviews in the USA, as we found per minute more competitive turn-taking events (χ^2 [1] = 5.9, $p = 0.01$) and more overlapping turns (χ^2 [1] = 5.2, $p = 0.02$), as well as longer turn overlaps ($F[1,310] = 15.6$, $p < 0.001$).

5.4 Discussion and Conclusion

According to Simone Weil, 'Humility is primarily a quality of attention' (1991) and despite many other philosophers had woven its praises, within political persuasion studies its persuasive effects have been neglected, focusing mainly on the expression of power. Recently D'Errico and Poggi (2019), starting from the definition of 'humble stance' in communicative terms, observed the bodily signals of humble politicians, showing how a leader's words, gestures, posture, facial expression can be cues of humility features like equality, informality, empathy. The aim of the present study is to explore features of humility associated with a speaker's tone-of-voice, by comparing two pairs of popular politicians from Usa (Obama vs Trump) and Italy (Gentiloni vs Salvini) during TV interviews. From results clear differences emerged between humble and dominant, involving prosodic characteristics from basically all fields of conversation behavior and tone of voice. Mainly three things distinguish the two types of politicians in both countries/cultures: (1) dominant leaders show a higher level of emotionality/expressiveness or arousal in their speech, reflecting here in a higher pitch register and variability as well as in a faster speaking rate; (2) they use a voice quality (Hammarberg & Formant dispersion) that signals urgency (we need to act now!!), power (acoustically larger body size), angriness or righteous indignation (we all know I

am right and you/they are not!), and thrusting. (3) dominant politicians also behave dominant/aggressive in a conversation, they talk rather than listen to what others (the TV host) have to say. Humble politicians talk without major interruptions, they rather repair their sentences and start over again (even multiple times) before they dare to stop talking and think before speaking, as the resulting silent pause could be used by the TV host to take the initiative. If the TV host tries to take the initiative they often simply continue talking (a lot louder also), sometimes even shouting the TV host down until he stop stalking.

These tone-of-voice characteristics and conversation behavior were qualitatively the same across the two countries, while from the quantitative point of view: the cultural differences can be reported mainly on Speaking & Pausing (i.e. silent pauses) and Speech melody parameters (i.e. pitch level), while other parameters need to be tested also in other speaking contexts.

More in general, in light of these first results we can discuss which of the parameters identified as characteristic of humble politicians are associated with the ones observed in charismatic leaders. According to these results, there are characteristics of humble politicians which are similar to those analyzed as charismatic (Niebuhr et al. 2019): the shorter IPU durations, the frequent silent pauses, and the higher formant dispersion is what charismatic speakers have in common with humble speakers. From another point of view the faster speaking rate, the larger pitch variation, and the higher levels of mean pitch and the Hammarberg index, for example, is what charismatic (as opposed to non-/less charismatic) speakers share with the dominant speakers. So, in other words, the charismatic speakers combine the urgent, powerful tone-of voice and the high emotionality of dominant speakers, with the smaller acoustic body size, positiveness, and interactional openness of humble speakers. On top come some perhaps charisma-specific parameter settings/changes, such as an extended pitch range and lower rather than higher pitch minima. Altogether, this shows that charismatic speech is based on a separate phonetic-prosodic profile that differs from "dominant" and "humble", but in parts overlaps with these two other profiles. Nevertheless, how strong this overlap actually is (for other parameters not analyzed here, such as emphatic-accent frequency, loudness level and variation, rhyhtmic variability, reduction level in consonant and vowel articulation etc.) still needs to be determined in future studies.

References

Boersma, P.: Praat, a system for doing phonetics by computer. Glot Int. **5**(9/10), 341–345 (2001)

Castelfranchi, C.: Micro-macro constitution of power. ProtoSociology Int. J. Interdiscip. Res. **18–19**, 208–265 (2003)

De Looze, C., Hirst, D.J.: Detecting changes in key and range for the automatic modelling and coding of intonation. In: Speech Prosody 2008, Campinas, Brazil (2008)

De Jong, N.H., Wempe, T.: Praat script to detect syllable nuclei and measure speech rate automatically. Behav. Res. Methods **41**(2), 385–390 (2009)

D'Errico, F., Poggi, I.: Tracking a leader's humility and its emotions from body, face and voice. In: Web Intelligence, vol. 17, no. 1, pp. 63–74. IOS Press (2019)

D'Errico, F., Poggi, I.: "Humble" politicians and their multimodal communication. In: Gervasi, O., et al. (eds.) ICCSA 2017. LNCS, vol. 10406, pp. 705–717. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-62398-6_50

D'Errico, F.: Humility based persuasion: Individual differences in elicited emotions and politician evaluation. (forth.)

D'Errico, F., Signorello, R., Demolin, D., Poggi, I.: The perception of charisma from voice: a cross-cultural study. In: 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII), pp. 552–557. IEEE (2013)

D'Errico, F., Poggi, I.: Blame the opponent! Effects of multimodal discrediting moves in public debates. Cogn. Comput. **4**(4), 460–476 (2012a)

D'Errico, F., Poggi, I., Vincze, L.: Discrediting signals. A model of social evaluation to study discrediting moves in political debates. J. Multimodal User Interfaces **6**(3–4), 163–178 (2012b)

Fitch, W.T.: Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. J. Acoust. Soc. Am. **102**, 1213–1222 (1997)

Fogg, B.J.: Persuasive technology: using computers to change what we think and do. Ubiquity 5 (2002)

Franzoni, V., Milani, A., Nardi, D.: Emotional machines: the next revolution. In: Web Intelligence, vol. 17, no. 1, pp. 1–7. IOS Press (2019)

Hammarberg, B., Fritzell, B., Gauffin, J., Sundberg, J., Wedin, L.: Perceptual and acoustic correlates of abnormal voice qualities. Acta Otolaryngol. **90**, 441–451 (1980)

Heldner, M.: Detection thresholds for gaps, overlaps and nogap-no-overlaps. J. Acoust. Soc. Am. **130**, 508–513 (2011)

Liu, X., Xu, Y.: Body size projection and its relation to emotional speech—evidence from Mandarin Chinese. In: Proceedings of Speech Prosody 2014, Dublin, pp. 974–977 (2014)

Niebuhr, O., Voße, J., Brem, A.: What makes a charismatic speaker? A computer-based acoustic-prosodic analysis of Steve Jobs tone of voice. Comput. Hum. Behav. **64**, 366–382 (2016). https://doi.org/10.1016/j.chb.2016.06.059

Niebuhr, O., Ochoa, S.G.: Do sound segments contribute to sounding charismatic? Evidence from a case study of Steve Jobs' and Mark Zuckerberg's vowel spaces. Int. J. Acoust. Vib. (2019, in press)

Nordström, H., Laukka, P., Thingujam, N.S., Schubert, E., Elfenbein, H.A.: Emotion appraisal dimensions inferred from vocal expressions are consistent across cultures: a comparison between Australia and India. R. Soc. open sci. **4**, 170912 (2017). https://doi.org/10.1098/rsos.170912

Poggi, I.: The "seeds" of charisma: Multimodal rhetoric of Mussolini's discourse. In: Tseronis, A., Forceville, C. (eds.) Multimodal Argumentation and Rhetoric in Media Genres, pp. 263–290. John Benjamins, Amsterdam (2017)

Poggi, I.: Mind, Hands, Face and Body. A Goal and Belief View of Multimodal Communication, pp. 1–432. Weidler (2007)

Poggi, I., D'Errico, F.: Dominance signals in debates. In: Salah, A.A., Gevers, T., Sebe, N., Vinciarelli, A. (eds.) HBU 2010. LNCS, vol. 6219, pp. 163–174. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-14715-9_16

Signorello, R., D'Errico, F., Poggi, I., Demolin, D.: How charisma is perceived from speech. A multidimensional approach. In: Proceedings of IEEE Social Computation, Amsterdam, 3–5 September, pp. 435–440. IEEE Computer Society (2012)

Vincze, L., Poggi, I., D'Errico, F.: Vagueness and dreams: analysis of body signals in vague dream telling. In: Salah, A.A., Ruiz-del-Solar, J., Meriçli, Ç., Oudeyer, P.Y. (eds.) HBU 2012. LNCS, vol. 7559, pp. 77–89. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-34014-7_7

Walker, M.B., Trimboli, C.: Smooth transitions in conversational interactions. J. Soc. Psychol. **117**, 305–306 (1982)

Weber, M.: The Theory of Social and Economic Organization. Oxford University Press, New York, USA (1920)

Weidman, A.C., Cheng, J.T., Tracy, J.L.: The psychological structure of humility. J. Pers. Soc. Psychol. **114**(1), 153 (2018)

Weilhammer, K., Rabold, S.: Durational aspects in turn taking. In: Proceedings of the International Conference of Phonetic Sciences (2003)

Xu, Y.: ProsodyPro — a tool for large-scale systematic prosody analysis. In: Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013), Aix-en-Provence, France, pp. 7–10 (2013)

Xu, Y., Lee, A., Wu, W.-L., Liu, X., Birkholz, P.: Human vocal attractiveness as signaled by body size projection. PLoS ONE **8**(4), e623 (2013)