Charisma perception in political speech: a case study

Rosario SIGNORELLO1,2, Francesca D’EERRICO2, Isabella POGGI2, Didier DEMOLIN1, Paolo MAIRANO1

1GIPSA-Lab, UMR 5216 CNRS, Grenoble INP, UJF, U. Stendhal (Grenoble, France); 2Dipartimento di Scienze dell’Educazione, Universita’ degli Studi Roma Tre (Rome, Italy)

rosario.signorello@gipsa-lab.fr, didier.demolin@gipsa-lab.fr, fderrico@uniroma3.it, poggi@uniroma3.it, paolomairano@gmail.com

Abstract

The charisma of the leader is conveyed through multiple aspects: his ideas and vision and his perceivable verbal and non-verbal behaviors. Among these perceivable behaviors there are the acoustic characteristics of speech. We present here a study on the perception of charisma in political speech. We collected speech statements with different illocutionary value taken from two speeches given by Umberto Bossi, the leader of an Italian party, before and after a stroke which caused him a voice disorder. Stimuli from the two condition differed significantly in the acoustic-prosodic features. In the first part of the study 40 French listeners rated normal speech stimuli (20 pre- and 20 post-stroke) and in the second part 22 French (11 pre- and 11 post-stroke) and 31 Italians (15 pre- and 16 post-stroke) rated the de-lexicalized version of the same stimuli. Results for the first part of the study show that pitch contour in Bossi’s pre-stroke speech positively influence the perception of his speech as charismatic, as opposed to those some years after the stroke. Results for the de-lexicalized speech confirm for French listeners our hypothesis of the influence of the pitch contour in Bossi’s charisma perception but they are controversial for Italian participants that seem to perceive Bossi as more charismatic in the post-stroke condition.

Keywords: charisma; political speech; intonation; illocution; voice disorder; speech synthesis.

1. Introduction

Charisma was firstly described by Weber as an “extraordinary quality” of a person who is believed to be endowed with superhuman properties thanks to which s/he gets ac-nowledged as a leader Cavalli, 1995: 5). Though no specific objective description of the “extraordinary quality” was given in Weber’s studies, some works started to study the perceivable behaviors of charismatic leaders: some, e.g., (Boss, 1976), focus on what we called the “charisma of the mind” (Signorello et al., 2012), that dwells in the strength of a leader’s ideas, others, e.g., (Atkinson, 1984) try to find visually or acoustically perceivable aspects of a leader’s behaviors that we called “charisma of the body” (Signorello et al., 2012). We suggest that both aspects of charisma, either jointly or independently, are responsible for its conveying and perception.

In the present study we focus on one aspect of the charisma of the body: the speech. We assume here that some of the perceivable acoustic-prosodic characteristics of a leaders speech are specifically responsible for conveying charisma. Our general goal is to characterize acoustically and distinguish perceptually a charismatic speech from a non-charismatic one.

Within previous work investigating the relationship between the acoustic-prosodic characteristics of a political leader’s speech and the perception of his/her charisma, Rosenberg and Hirschberg, 2009) studied the correlation between acoustic, prosodic, and lexico-syntactic characteristics of political speech and the perception of charisma; Touati (1993) investigated the prosodic features of rhetoric utterances in French political speech in pre and post-elections discourses. Other works examined the relationship between prosodic features and the perception of a speaker as a “good communicator” (Strangert & Gustafson, 2008) or analyzed the pitch contour of French political leaders’ speech and its idiosyncratic and contextual variations (Martin, 2009).

2. A hypothesis about charisma

According to Poggi (2005), in persuasive discourse the speaker tries to convince the audience to do some action by exploiting the three strategies posited by Aristotle (2011): Logos (the rational argument), Pathos (the appeal to the audiences emotions), and Ethos (the character of the speaker). According to the theory of Poggi (2005) and Poggi et al. (2011), the dimension of Ethos also includes, for the political leader, three sub-dimensions: Benevolence (the tendency to act in the interest of the audience), Competence (the capacity for rational foreseeing and planning), and Dominance (the power to prevail in a competition).

The notion of charisma we proposed in (Signorello et al., 2012) is based on this theoretical framework. We defined charisma as a set of characteristics of a leader that include his “having a vision” (a goal towards which he wants to lead his followers), a “high level of dominance” (look strong, persistent and fighting) and “emotional intelligence” (the ability to feel and transmit emotions, and to be and look empathic). The combination of these features makes a leader charismatic, and is displayed by his/her non-communicative and communicative behavior.

3. What makes a speech charismatic?

To investigate the perception of charisma in political speech we analyzed the acoustic and prosodic characteristics in the speech of Umberto Bossi, an Italian politician who in 2004, during his political career, had a stroke that resulted in severe speech impairment. We collected two samples taken from two speeches performed, respectively, in 1994 (the pre-stroke condition, PRE) and in 2011 (the post-stroke condition, POST). Our hypothesis was that the important differences in
acoustic-prosodic characteristics of Bossi’s speech, in samples of political speeches preceding and following the stroke, give rise to a different perception of charisma. If this hypothesis is validated perceptually we might conclude that information about charismatic qualities are borne by the acoustic-prosodic characteristics that differ in the two samples.

In order to describe the charisma phenomenon through common language adjectives we conducted a qualitative study collecting adjectives describing what charisma is and what it is not (a brief summary is presented in section 3.1.. For the extensive study see (Signorello et al., 2012). We then analysed Bossi’s acoustic-prosodic features in the PRE and POST and conducted a language-independent perceptual study on French participants (section 3.2.4.). We then de-lexicalized our stimuli by synthesis only preserving the pitch contour, the duration and the intensity and conducted a perceptual study on French and Italian listeners. In isolating the pitch contour we could verify if this is the aspect that influences the perception of charisma in Bossi’s speech (section 3.3.).

3.1 Describing charisma

In a previous work (Signorello et al., 2012) we constructed a questionnaire aimed to assess the perception of charisma in the samples of Bossi’s speech required to previously make up a list of adjectives that express charismatic and non-charismatic qualities. To find out such adjectives in an empirically grounded way, we administered a questionnaire through Internet to 58 French participants (42 female, 16 male, mean age 30), asking to freely generate adjectives connected to the idea of what charisma is and what it is not. We obtained a list of French adjectives, 106 describing charisma positively and 105 describing what charisma is not. In order to make a manageable questionnaire, we further selected 67 adjectives (Table 1) retaining only those occurring more than once, 42 positively and 20 negatively related with charisma. We then classified those adjectives in a multidimensional scale of charisma under five dimensions describing this phenomenon. An extended report of this multidimensional scale of charisma and on how adjectives describing charisma are classified in it can be found in (Signorello et al., 2012).

3.2 Normal Speech

3.2.1. Stimuli

Previous works about the perception of a speaker as a good (Strangert & Gustafson, 2008) or charismatic speaker (Rosenberg & Hirschberg, 2009) rely on the acoustic analysis and the perceptual evaluation of stimuli classified per speaker, topic and genre of speech. Our approach is different. We chose 3 stimuli per condition (PRE and POST) according to their illocutionary value: an assertion, an in-citation and a rhetorical wh-question. As we know the speaker shapes prosody differently in relation to different speech acts (Firenzueli, 2001). Our hypothesis is that all three types of speech acts are perceived as more charismatic in the PRE condition thanks to prosodic features. Further we argue that incitation might be perceived as more charismatic than rhetorical question which in turn might be perceived as more charismatic than assertion. Below we describe the acoustic-prosodic features of our stimuli.

<table>
<thead>
<tr>
<th>DIMENSION</th>
<th>PRE</th>
<th>POST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pathos</td>
<td>passionate, empathetic, enthusiastic, reassuring</td>
<td>cold, indifferent</td>
</tr>
<tr>
<td>Ethos Benevolence</td>
<td>extraverted, positive, spontaneous, trustworthy, honest, fair, friendly, easygoing, makes the others feel important</td>
<td>untrustworthy dishonest, egocentric, individualistic, introverted</td>
</tr>
<tr>
<td>Ethos Competence</td>
<td>visionary, organized, smart, sagacious, creative, competent, wise, enterprising, determined, resolute, who propose, seductive, exuberant, sincere, clear, communicative</td>
<td>inefficient, inadequate, uncertain, faithless, unclear, menacing</td>
</tr>
<tr>
<td>Ethos Dominance</td>
<td>dynamic, calm, active, courageous, confident, vigorous, strong, leader, authoritarian, captivating, who persuade, who convince</td>
<td>apathetic, timorous, weak, conformist, unimportant, who scare</td>
</tr>
<tr>
<td>Emotional Induction Effects</td>
<td>charming, attractive, pleasant, sexy, bewitching, eloquent, influential</td>
<td>boring</td>
</tr>
</tbody>
</table>

Table 1: The 67 positive and negative adjectives related with charisma collected among the naïve French participants (in English for clarity purposes). Reprinted from Signorello et al., 2012)

3.2.2. Overall F0 measures

The PRE speech presents higher F0 means than the POST speech: PRE (F0 mean 178.89 Hz; min 101.84 Hz; max 241.10 Hz), POST (F0 mean 120.20 Hz; min 91.78 Hz; max 155.99 Hz). All means from the PRE differ significantly from the POST (p<0.0001). Our findings confirm and extend (Murry, 1978)’s findings on significant differences in F0 measures between normal and disordered voice. We argue that F0 values might be positively correlated to charisma perception.

3.2.3. Pitch contour description

The assertion in the PRE condition (Figure 1a below) presents a syntactic focus on “questo” [this], emphasized by a high fall and separated by a pause from the rest of the sentence. The right-side part of the tonal unit presents a falling contour with a small peak on the last tonic syllable. Instead, in the POST condition (Figure 2a below) the
sentence presents a moderate falling and flat pitch contour with a peak on the third lexical word. The incitation in the PRE condition (Figure 1b) includes two parts, each with a perception of the acoustic signal was optimal and that the semantic content was not understood. Then they had to express their judgment about the stimuli through our 67-adjective inventory on a 7-point Likert scale pitch contour starting with high frequency and falling sharply in the last tonic syllable. In the POST condition instead the incitation (Figure 2b) presents two rising-falling contours in the first part and falls gradually in the right part of the tonal unit. The rhetorical wh-question in the PRE condition (Figure 1c) presents two contiguous pitch contour movements: the rising part corresponds to the wh-element, the falling part corresponds to the verb. A gradual falling movement comes on the right side of the tonal unit. In the POST statement (Figure 2c) a falling contour corresponds to the wh-element and a rising contour to the verbal element, with a gradual falling movement on the right side of the tonal unit.

3.2.4. Perception experiment

Forty French participants with no knowledge of Italian rated the stimuli presented in the section above via an HTML/PHP browser-based interface. Twenty of them listened to the PRE condition and twenty to the POST condition stimuli. The test took place in an anechoic chamber and participants wore a Sennheiser HD 25-13 headphone. After listening to each stimulus a participant had to answer to some check questions to verify that the

Figure 1: Intonation contour, transcription, translation, du-ration and F0 measures of PRE stimuli per speech act. (a): Assertion. “Questo amici ereditiamo” [This, my friends, is what we inherit]. 3.51s. F0 mean 52.62 Hz; SD 12.40 Hz; min 95.25 Hz; max 210.94 Hz; range 13 ST. (b): Incitation. “Si ritorna all’attacco, fuori dalle trincee” [Let’s take up again the offensive, get out of the trenches]. 4.27s. F0 mean 225.51 Hz; SD 38.58 Hz; min 107.74 Hz; max 270.36 Hz; range 16 ST. (c): Rhetorical wh-question. “E come facevamo a farlo?” [How could we have done it?]. 1.81s. F0 mean 138.28 Hz; SD 27.98 Hz; min 96.07 Hz; max 189.39 Hz; range 11.72 ST. Spectrogram and pitch contour graphics obtained with WinPitch software (Martin, 2011)

Figure 2: Intonation contour, transcription, translation, du-ration and F0 measures of POST stimuli per speech act. (a): Assertion. “Noi siamo schiavi del centralismo romano” [We are slaves of the Roman centralism]. 2.46 s, F0 mean 116.77 Hz, SD 10.74 Hz, min 86.64 Hz, max 146.45 Hz, range 9 ST. (b): Incitation. “La Lega è pronto per conquistare la libertà della padania” [The Lega is ready to conquer the freedom of padania]. 6.61s, F0 mean 142.02 Hz, SD 38.58 Hz, min 86.2 Hz, max 182.08 Hz, range 12 ST. (c): Rhetorical wh-question “E come fanno a lavorare questa gente?” [How can these people work?]. 1.89 s, F0 mean 117.93 Hz, SD 15.54 Hz, min 90.56 Hz, max 192.99 Hz, range 13 ST. Spectrogram and pitch contour graphics obtained with WinPitch software (Martin, 2011)
to Ethos Benevolence results are quite inconsistent: the adjectives attributed to the PRE speech include egocentric, dishonest and individualistic, which in our previous qualitative study (Table 1) are non-charismatic qualities.

### Table 2: Adjectives describing the perception of charisma in the Bossi’s speech by condition with rating values (t-test, p<.001)

<table>
<thead>
<tr>
<th>ADJECTIVES</th>
<th>PRE</th>
<th>POST</th>
</tr>
</thead>
<tbody>
<tr>
<td>dynamic</td>
<td>5.48</td>
<td>4.12</td>
</tr>
<tr>
<td>authoritative</td>
<td>6.19</td>
<td>4.57</td>
</tr>
<tr>
<td>calm</td>
<td>2.66</td>
<td>1.76</td>
</tr>
<tr>
<td>extraverted</td>
<td>1.66</td>
<td>1.23</td>
</tr>
<tr>
<td>timorous</td>
<td>3.03</td>
<td>3.99</td>
</tr>
<tr>
<td>wise</td>
<td>2.85</td>
<td>2.28</td>
</tr>
<tr>
<td>individualistic</td>
<td>4.81</td>
<td>4.61</td>
</tr>
<tr>
<td>active</td>
<td>4.19</td>
<td>4.28</td>
</tr>
<tr>
<td>introverted</td>
<td>1.52</td>
<td>1.23</td>
</tr>
<tr>
<td>menacing</td>
<td>4.57</td>
<td>5.33</td>
</tr>
<tr>
<td>energetic</td>
<td>5.14</td>
<td>6.09</td>
</tr>
</tbody>
</table>

Table 3: Adjectives describing the perception of charisma in the Bossi’s speech by association, incitation and rhetorical wh-question and condition with rating values and one-way ANOVA’s values (p<.001). Higher rates in bold

As for the dimensions of Ethos Competence and Ethos Dominance our hypothesis is almost completely validated: the speaker is perceived as competent, smart, clear, seductive, etc. in the PRE and as unclear in the POST; as dynamic, authoritarian, confident, leader in the PRE and as boring in the POST speech. These results validate our hypothesis on the attribution of charismatic qualities to the PRE as opposed to the POST speech.

Taking into account the different types of speech act both in the PRE and in the POST speech the different illocutionary act elicits a different perception. The incitation is the one that influences the most the perception of charisma. In particular for the dimension of Ethos Competence the incitation elicits adjectives as competent (F(2, 123)=3.114; p<0.048), resolute (F(2, 123)=6.767; p<0.002), enterprising (F(2, 123)=8.515; p<0.001), clear (F(2, 123)=3.046; p<0.05), exuberant (F(2, 123)=4.23; p<0.017) and communicative (F(2, 123)=2.705; p<0.05). More than other speech acts the incitation has a significant effect on the perception of the speaker’s emotional state (see adjectives as passionate (F(2, 123)=2.999; p<0.05), influential (F(2, 123)=9.359; p<0.001) and enthusiastic (F(2, 123)=4.765; p<0.01)).

The assertion on the other hand evokes more non-charismatic qualities like indifferent (F(2, 123)=3.459; p<0.035) and unclear (F(2, 123)=3.626; p<0.029). Finally the rhetorical question seems to not influence a specific dimension of charisma. However, if we consider effect of both condition and a particular speech act the results are quite different. Through a one-way ANOVA we crossed the results of the condition (PRE vs. POST) and the different types of speech act (assertion, incitation and rhetorical wh-question) to study the influence of the different illocutionary acts on the perception of Bossi’s charisma (see Table 3). The incitation makes Bossi to be perceived as more dynamic, authoritarian, active, menacing, and energetic in the PRE condition and as extraverted and wise in the POST condition. Through the Assertion he has been perceived as individualistic in the PRE speech and as calm and introverted in the POST speech. As for the rhetorical wh-question the only significantly results is timorous in the PRE speech.

### 3.3 Synthesized speech

#### 3.3.1 Stimuli

We decided to carry out a perceptive test on de-lexicalized stimuli in order to further validate our hypothesis that the pitch contour is a relevant element for the perception of charisma. In fact, our de-lexicalization procedure enables us to isolate the pitch contour of a sentence from the semantic content, segmental features and voice quality characteristics. In this way, listeners are therefore forced to give their judgments solely on the basis of intonation, all other linguistic information being eliminated. The de-lexicalized procedure we chose has been developed for the AMPER (Atlas Multimédia Prosodique de l’Espace Roman) project developed by Albert Rilliar on the basis of scripts originally elaborated by Antonio Romano (see Contini et al., 2002 for details). It consists in synthesizing a periodic waveform with the original pitch, intensity and duration values of the actual sentence (this is done by taking three measures per vowel, respectively at the onset,
middle and offset-consonants are replaced with silence). This procedure has been used by several authors working on the AMPER project and has already proved its efficacy.

3.3.2. Perception experiment
Twenty-two French (11 PRE, 11 POST) and thirty-one Italian (15 PRE, 16 POST) listeners participated to a perception analysis with the same methodology described in section 3.2.4. Thus the only differences were the de-lexicalized stimuli.

3.3.3. Results
The first results for the de-lexicalized stimuli perception, compared to results for normal speech perceptions, confirm in one hand our hypothesis of the influence of the pitch contour in Bossi’s charisma perception for French participants but they are, in the other hand, controversial for Italian participants. In fact French listeners describe Bossi as charming, who propose, timorous, confident, pleasant, introverted in the PRE speech and as inadequate, spontaneous, active, leader in the POST speech (t-test, p<0.05). For Italian participants we only performed the perceptual test of de-lexicalized stimuli in order to avoid semantic and ideology influence on the perception of Bossi’s speech. Italian listeners perceived the speaker as boring, indifferent and unimportant in the PRE speech and as attractive, visionary, sexy, cold, passionate, seductive in the POST speech (t-test, p<0.05). From these preliminary results it seems that the pitch contour-only stimuli elicit a different type of of Bossi’s charisma for Italians listeners. In fact the POST speech is described with adjectives positively related with charisma and the PRE speech with adjectives describing charisma negatively, a trend in results that goes against our theory of pre-stroke speech as more charismatic than the post-stroke.

4. Conclusion
In this study we aimed to demonstrate that the perception of charisma in political speech is partly determined by the acoustic characteristics of speech. To do so, we first analyzed samples from the speech of the Italian politician Umberto Bossi before and after a stroke; through a qualitative study we singled out 67 adjectives describing charismatic and non-charismatic qualities. Finally we run a perception study asking participants to rate Bossi’s samples in terms of those adjectives. As resulted from the acoustic analysis, the PRE speech, with its intonation features as focus words, tonal jumps, and higher values, dramatically differs from the POST. And since the results of the perception study validate our hypothesis that Bossi’s speech after the stroke is perceived as less charismatic than before, we may reasonably conclude that the characteristics of intonation that differentiate Bossi’s PRE and POST speeches are an important factor in the perception of charisma. And this hypothesis has been validated once more through a perceptual experiment in which we only tested the intonation contour influence on the Bossi’s charisma perception. We also de-lexicalized stimuli and preserved original pitch, intensity and duration values and we tested French and Italian participants. Results validate our hypothesis on the intonation contour relevance on charisma perception of the PRE speech for French participants but are controversial for Italians. In any case our results on synthesized speech are preliminary and they will be statistically analyzed more in depth. Naturally we are aware that the acoustic characteristics of speech also include voice quality, which we think is relevant too. In future work we will investigate the importance of voice quality in determining the perception of charisma, while trying to distinguish it from the contribution of intonation, also through synthesis of speech fragments.

5. Acknowledgements
We would like to thank Christophe Savariaux (GIPSA-Lab, CNRS, Grenoble, France) for his technical support in acoustic analysis. Research partly supported by a Doctoral Fellowship from the French Ministry of National Education and Research (years 2010-2013) and by the 7th Framework Program, European Network of Excellence SSPNet (Social Signal Processing Network), Grant Agreement N. 231287.

6. Acknowledgements


