The physiological use of the charismatic voice in political speech

Rosario Signorello[1,2], *Didier Demolin*[1]

[1]Grenoble Alps University, GIPSA-Lab, Grenoble, France
[2]Educational Department, Roma Tre University, Rome, Italy
{rosario.signorello, didier.demolin}@gipsa-lab.fr

Abstract

Voice is one of the orators perceivable behaviors that convey traits of charisma. Contemporary leaders, such as politicians, use their voices to shape their message and persuade audiences. The present study analyzes the physiological use of the "charismatic voice" in Italian, French, and Brazilian Portuguese political speech. We investigated the "discourse level" aspects of voice by conducting acoustic analyses on the voice production of political leaders speaking in different communication contexts (formal versus informal) and while addressing different kinds of audiences (interviewers, colleagues, voters). Politicians' physiological speech range profiles were found to be dependent on the communication context, the audience being addressed and the persuasive goal. There thus seems to be a cross-language and cross-cultural trend in common for voice usage in different communication contexts. Orators adapt voice behavior to the context, the audience, and the persuasive goal of communication. The range of acoustic voice correlates will be wider when the discourse must be at its most persuasive.

Index Terms: Physiology of voice, charisma, political speech, speech range profiles.

1. Introduction

Charisma has been described in various manners: as the quality of leaders that are capable of combing vision and mental and physical skills to preside over others (Socrates 469–399 BC); as the "extraordinary quality" of a person who is believed to be endowed with superhuman properties that make him stand out as a leader [1]; or as a socially constructed phenomenon where in leaders act as the entrepreneurs of their identity in order to establish a relationship with followers [2]. We define charisma as a set of leadership characteristics that are displayed through a persons perceived non-verbal communication behaviors (the "charisma of the body") that convey the strength of his ideas and vision through verbal behavior and written texts (the "charisma of mind") [3]. We also suggest that both of these two aspects of charisma, either jointly or independently, are responsible for conveying charisma. The voice is one of the perceivable behaviors of the "charisma of the body" and its acoustic parameters are responsible for conveying charisma. In previous works (e.g., [4]) we distinguished a charismatic voice from a non-charismatic one through samples of political speech. We also showed how different acoustic voice patterns convey different charisma traits and how they could influence perception in the process of persuasion in political discourse [3].

In section 2 we describe our multi-lingual and multi-cultural corpora of political speeches (in Italian, French and Brazilian Portuguese) and we analyze the voice of political speech through speech range profiles. Then, in section 3 we investigate the physiological acoustic correlates of political speech by analyzing the acoustic correlates of orators' voices in different communication contexts (formal and informal discourse, addressing different kinds of audiences). Acoustic voice parameters were plotted onto charts in order to quantify the physiological voice range during the speeches and to verify the distinctive use of ones voice throughout the various communication contexts. We then carried out a cross-language and cross-cultural comparison of these results which will be discussed in the same section in order to emphasize the similarities and differences of voice usage in political speech. In section 4, we discuss the time-related adjustment of voice acoustic correlates in different communication contexts. Studying voice adjustment throughout the entire speech allowed us to exhaustively investigate the entire voice range used and to explain how voice is used to pursue a persuasive versus a non-persuasive goal.

2. Method

2.1. Material

We collected corpora for 3 speakers: Luigi de Magistris (S-IT), the current mayor of Naples, Italy; François Hollande (S-FR), the current President of the French Republic; and Luiz Incio Lula da Silva (S-BP), the former President of Brazil. For each of these speakers we collected speeches given in three different communication contexts: an "informal interview addressed to the interviewer" (INT); a "formal conference addressed to other politicians" (CON); and a "formal monologue addressed to followers" (MON). The recordings were all retrieved from various public archives available online[1]. Duration of recordings: S-IT: INT –12m12s–; CON –06m49s–; MON –14m56s–. For S-FR: INT –24m14s–; CON –11m14s–; MON –1h25m45s–. For S-BP: INT –4m43s–, CON –18m53s–, MON –29m51s–. All of the recordings include the entire discourse, which allows us to analyze the entire vocal range during a speech.

2.2. Speech Range Profile (SRP)

To quantify the minimum and maximum physiological voice limits of speakers throughout different communication contexts (INT, CON, and MON), we used Speech Range Profiles (SRP). SRPs are defined as physiological voice range profiles recordings that are specifically designed for recording continuous speech [5, p. 26]. In this article, we refer to SRP as the chart of dynamic range and the minimum and maximum averages of acoustic voice parameters in spontaneous political speech. As showed in Figures 1, 2 and 3, SRPs are two-dimensional charts plotted with F0 and Intensity average values computed from /a/ vowels[2] uttered by speakers from our corpora. The F0 scale (X-

[1]Sources: www.rai.tv - www.elysee.fr - www.institutolula.org

[2]/a/ vowels have a high first formant (Catford [6, p. 154] reports an average frequency of about 850 Hz for the F1 of [a] vowels, in a

axis) ranges from 75 to 350 Hz (the average range of male adult speakers) and the Intensity scale (Y-axis) ranges from 40 to 120 dB (as recommended by The Union of European Phoniatricians [9, p. 117]). SRPs were automatically plotted using R [8].

As previously stated by Lamarche [5, p. 21] mapping of the interaction between F0 and Intensity is useful to illustrate the speaker's voice range skills in controlling and varying these acoustic parameters. According to the literature these two parameters correlate: maximum and minimum Intensity values increase as F0 values rise [9, p. 111-112]. SRP mapping is therefore very helpful in this study because it reveals the differences in voice usage in a single language with respect to the communication context, and it highlights the similarities or differences among the political speech of different languages and cultures.

3. Acoustic interpretation of SRPs

3.1. Hypotheses

Politicians address people with the goal of persuasion [7]. In order to reach this goal, the speaker uses verbal and non-verbal communication to convince followers to accept his goal and help him to accomplish its fulfillment [10]. We argue that speakers use different voice acoustic range in different communication contexts as follows: in MON, where the main goal is to persuade followers to vote, the speaker "stretches" his voice to maximum and minimum acoustic limits. In CON, wherein the goal is to persuade peers, i.e. other politicians, the speakers acoustic voice range is less wide, whereas in INT, where the speaker does not deal with any political topic and the persuasive goal is different and/or non-existent, the speaker uses an even narrower acoustic voice profile. Through the use of acoustic analysis, SRPs, and statistical F0 and Intensity analysis, this section examines how orators modulate acoustic voice correlates in the three communication contexts in question.

3.2. Luigi de Magistris (S-IT)

In INT, S-IT's voice profile covers the low F0 region (in Hz: min 81, mean 130, max 200. Range 15 ST) and the low-Intensity region (in dB: min 51, mean 64, max 70) of the SPR (see Figure 1a). There is no significant correlation ($p>.05$) between F0 and Intensity values. During CON, his voice covers the middle/low F0 region (in Hz: min 79, mean 148, max 256. Range 20 ST) and the middle/low Intensity region (in dB: min 55, mean 68, max 77) of the SRP (see Figure 1b). We found significant positive correlation between the increasing of F0 and Intensity values ($t(258)=14.16$, $p<.001$, $r=.77$). During MON, the voice profile occupies the high/low F0 region (in Hz: min 81, mean 183, max 290. Range 22 ST) and the middle of the Intensity region (in dB: min 57, mean 74, max 88) of the SRP (see Figure 1c). We found significant positive correlation between the increasing of F0 and Intensity values ($t(853)=20.69$, $p<.001$, $r=.76$).

3.3. François Hollande (S-FR)

In INT, S-FR's voice profile covers the low-F0 region (in Hz: min 79, mean 110, max 261. Range 20 ST) and the middle/low-Intensity region (in dB: min 55, mean 68, max 81) of the SRP (see Figure 2a). We found significant positive correlation between the increasing of F0 and Intensity values ($t(371)=11.40$, $p<.001$, $r=.71$). During CON, his voice extension occupies the middle/low-F0 region (in Hz: min 85, mean 146, max 236. Range 17 ST) and the middle/low-Intensity region (in dB: min 55, mean 72, max 77) of the SRP (see Figure 2b). We found significant positive correlation between the increasing of F0 and Intensity values ($t(291)=11.91$, $p<.001$, $r=.75$). During MON, his voice extension occupies the high/low-F0 region (in Hz: min 76, mean 212, max 298. Range 24 ST) and the middle-Intensity region (in dB: min 53.32, mean 73.04, max 79.43) of the SRP (see Figure 2c). We found significant positive correlation between the increasing of F0 and Intensity values ($t(1566)=15.77$, $p<.001$, $r=.60$).

male adult voice), and this reduces the possibility that the F0 frequency measures will conflict with the F1 frequency [5, p. 26].

3.4. Luiz Inácio Lula da Silva (S-BP)

In INT, S-BP's voice profile covers the low-F0 region (in Hz: min 76, mean 100, max 156. Range 12 ST) and the middle Intensity region (in dB: min 55.76, mean 68.45, max 81.70) of the SRP (see Figure 3a). There is no significant correlation ($t(189)=7.33$, $p>.05$) between F0 and Intensity values. During CON, his voice extension occupies the middle/low-F0 region (in Hz: min 77, mean 141, max 238. Range 19 ST) and the low-Intensity region (in dB: min 46.05, mean 62.92, max 71.22) of the SRP (see Figure 3b). We found positive significant correlation between the increasing of F0 and Intensity values ($t(1126)=14.24$, $p<.001$, $r=.62$). During MON, his voice extension occupies the high/low-F0 region (in Hz: min 87, mean 176, max 292. Range 21 ST) and the middle/low-Intensity (in dB: min 53, mean 73, max 79) region of the SRP (see Figure 3c). We found positive significant correlation between the increasing of F0 and Intensity values ($t(787)=13.25$, $p<.001$, $r=.65$).

3.5. Discussion

In this section we analyzed our three speakers specific use of voice in different communication contexts. A Kruskal-Wallis test indicates that there is a significant difference in the F0 and Intensity means of the voice profiles between the three different speech contexts: S-IT (F0 $H(2)=532.73$, $p<.001$; Intensity $H(2)=697.14$, $p<.001$); S-FR (F0 $H(2)=1201.43$, $p<.001$; Intensity $H(2)=400.57$, $p<.001$); S-BP (F0 $H(2)=687.70$, $p<.001$; Intensity $H(2)=1329.94$, $p<.001$). SRPs acoustic parameters are therefore significantly different between the contexts of communication.

We then conducted *post hoc* multiple comparison tests (see [11, p.681-684]) in order to determine the differences among the three separate communication contexts for each of the speakers. F0 values in MONs were significantly higher than CON (S-IT: *diff.*=400.44; S-FR: *diff.*=843.65; S-BP: *diff.*=508.44) and INT (S-IT: *diff.*=644.11; S-FR: *diff.*=1154.25; S-BP: *diff.*=1179.99). F0 values in CON were significantly higher than INT (S-IT: *diff.*=243.66; S-FR: *diff.*=310.60; S-BP: *diff.*=671.65). In all cases, $\alpha=.05$. Intensity values in MONs were significantly higher than in CONs (S-IT: *diff.*=451.21; S-FR: *diff.*=216.05; S-BP: *diff.*=972.05) and than INTs (S-IT: *diff.*=742.05; S-FR: *diff.*=740.62), except for S-BP (*diff.*=14.17), does not present any note-worthy difference. Intensity values in CONs were found significantly higher than INT for S-IT (*diff.*=290.84) and S-FR (*diff.*=524.56). For S-BP in INT, Intensity values were significantly higher than in CON (*diff.*=957.87). In all cases $\alpha=.05$.

The specific comparisons of F0 and Intensity mean ranks between INT, CON, and MON presented above confirm what we see in SRPs: the context of communication significantly influences the voice range profile. We can conclude that the speaker adjusts his acoustic voice correlates depending on the

communication context. And for at least the speech samples analyzed, these results are cross-language and cross-cultural.

4. Time-related adjustment of voice

4.1. Hypotheses

Rhetorical speeches are usually organized in a sequence of propositions of increasing importance or force that build up to a climax at the end of the speech, in unison with the audiences response (a structure similar to the climax figure-of-speech but throughout the whole speech). In this section, we investigate how speakers adjust their acoustic voice correlates through the lapse of time in the three different communication contexts. We argue that in formal speech in which discourse should be more persuasive, there is a relationship be- tween the adjustment of acoustic voice correlates and time. We also argue that in non-rhetorical discourse, wherein the speaker does not deal with political topics and the persuasive goal (if any) is different than it is in formal discourse, there is no time-related adjustment. We used simple linear regression to show patterns in time-related adjustment acoustic voice correlates by fitting a model to our data and using it to predict the values of the dependent variables, F0 and Intensity average value of /a/ vowels, based on the independent variable of Time.

4.2. Luigi de Magistris (S-IT)

The regression model applied on S-IT's data shows a significant increasing of F0 during time for INT ($F(210)=8.46$, $p<.01$, $r=.19$), CON ($F(358)=19.02$, $p<.001$, $r=.22$), and MON ($F(210)=8.46$, $p<.001$, $r=.21$). The regression model also shows a significantly decreasing of Intensity during time in INT ($F(210)=13.93$, $p<.001$, $r=.24$) and MON ($F(853)=16.92$, $p<.001$, $r=.17$) and a significant increasing during CON ($F(358)=11.26$, $p<.001$, $r=.17$). S-IT shows a clear between-communication context trend in F0 values adjustment (F0 always increases) that we did not find for Intensity values (he decreases them during INT and MON while increasing them in CON).

4.3. François Hollande (S-FR)

The regression model applied on S-FR's data shows a significant decrease of F0 throughout time for MON ($F(1566)=14.13$, $p<.001$, $r=.09$) but not for INT and CON ($p>.05$). For the Intensity the regression model shows a significant decrease throughout time for MON ($F(1566)=22.1$, $p<.001$, $r=.12$), and CON ($F(291)=4.17$, $p<.05$, $r=.12$). In INT, the speaker makes no significant Intensity value adjustment as the time elapses ($p>.05$). S-FR shows no overall tendency of acoustic correlate adjustment throughout time. While he significantly decreases both F0 and Intensity during MON, he significantly decreases only Intensity in CON. INT data do not fit the model ($p>.05$).

4.4. Luiz Inácio Lula da Silva (S-BP)

The regression model applied on S-BP's data shows on one hand a significant time-dependent F0 increase during time for MON ($F(787)=63.13$, $p<.001$, $r=.27$) and CON ($F(1126)=21.21$, $p<.001$, $r=.13$), but a significant time-dependent F0 decrease during time for INT ($F(189)=7.64$, $p<.01$, $r=.19$) on the other hand. For the Intensity parameter, the regression model shows a significant increase throughout time in MON ($F(787)=12.92$, $p<.001$, $r=.12$), and no significant adjustment in INT and CON ($p>.05$).

4.5. Discussion

The results show how speakers adjust acoustic voice correlates during time in different communication contexts. We found some similar patterns (between certain subjects) but not any overall cross-language and cross-cultural trends. As for F0 adjustment through time, S-IT and S-BP speakers significantly increase their F0 in CON and MON, i.e. the formal contexts of communication. However, the speaker S-FR significantly decreases his F0 during MON (INT and CON are not significant). As for Intensity values, our model shows a significant decrease of Intensity during MON for S-IT and S-FR whereas it shows a significant increase for S-BP. During CON, S-IT significantly increases Intensity while S-FR significantly decreases it. S-BP data do not fit the model. And in INT, only S-ITs data fit the model: the speaker significantly decreases the Intensity as time lapses.

5. Conclusion and Perspectives

In this work we investigated the physiological voice productions in political speech across three languages and cultures, throughout different communication. Results from acoustic and statistical analyses show a similar cross-language and cross-cultural use of ones voice during political discourse. Speech range profiles results were significantly dependent on the communication context. Orators from all languages use a small range of acoustic voice correlates in informal discourse in which the speaker does not deal with political topics and the persuasive goal is different from what it might be in formal discourse. In rhetorical formal discourse in which the speaker must be more persuasive, the voice range profile widens. And while addressing peers, i.e. other politicians, the acoustic voice acoustic correlates cover a larger range of minimum and maximum values.

We also investigated the adjustment of acoustic voice correlates throughout time. Orators from our corpora significantly adjust their voice correlates as time elapses during a discourse. There are no overall cross-language and cross-cultural similarities in our speech sample that could be used to theorize this phenomenon. Orators therefore assume different vocal behaviors that are more related to the individual speaker than to their language/culture.

Several previous studies on voice production and/or perception have pointed out how ones voice could convey personality and charisma traits (for example, [12] and [13] for personality; [14] or [3] for charisma traits). The speaker adjusts his acoustic voice correlates in order to appear a certain way (primary meaning) and to achieve a goal (secondary meaning) (see [15] on voice and [10] on persuasion). We have presented a study here on the voice production of modern society leaders, i.e. politicians. Leaders convey their personality and charisma traits through their voice, among other behaviors, when they address an audience. They aim to fulfill the "primary meaning" of appearing to be a charismatic leader and the "secondary meaning" of using their own behaviors to be persuasive and actually became the leader.

Our study is the first step in attempting to show how a speaker physiologically shapes his voice behavior to be perceived as charismatic so that he will be recognized as a leader. Our findings about similarities in voice production across languages and cultures (at least for political speech and the sample of language we studied) will lead us to further investigate the "culture-specific and learned" versus the "biological basis" of speech production.

(a) INT (b) CON (c) MON

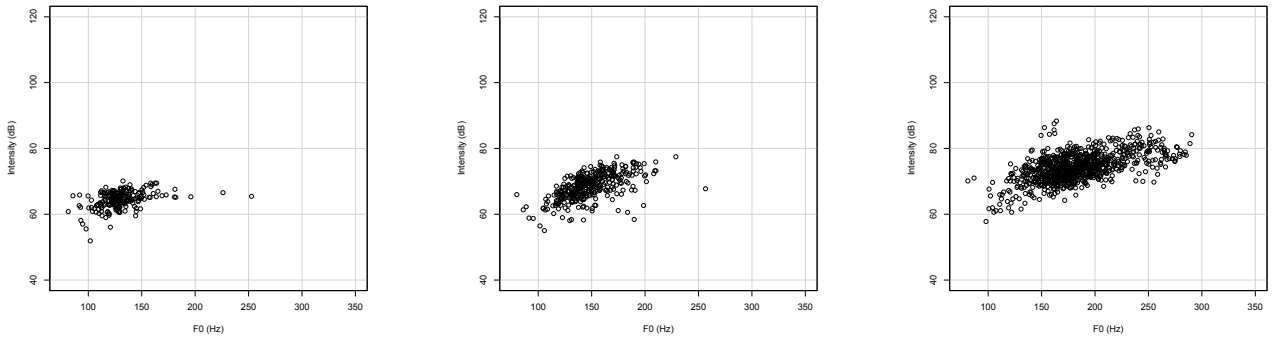

Figure 1: Speech Range Profiles of Italian speaker Luigi de Magistris (S-IT). Communication context: (a) informal interview addressed to the interviewer (INT); (b) formal conference addressed to other politicians (CON); (c) formal monologue addressed to followers (MON). X-axis: F0 (Hz), Y-axis: Intensity (dB).

(a) INT (b) CON (c) MON

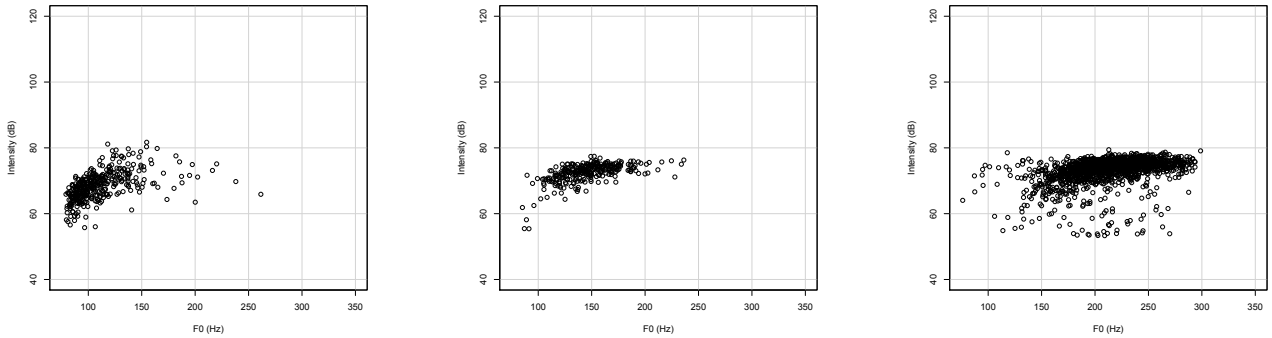

Figure 2: Speech Range Profiles of French speaker François Hollande (S-FR). Communication context: (a) informal interview addressed to the interviewer (INT); (b) formal conference addressed to other politicians (CON); (c) formal monologue addressed to followers (MON). X-axis: F0 (Hz), Y-axis: Intensity (dB).

(a) INT (b) CON (c) MON

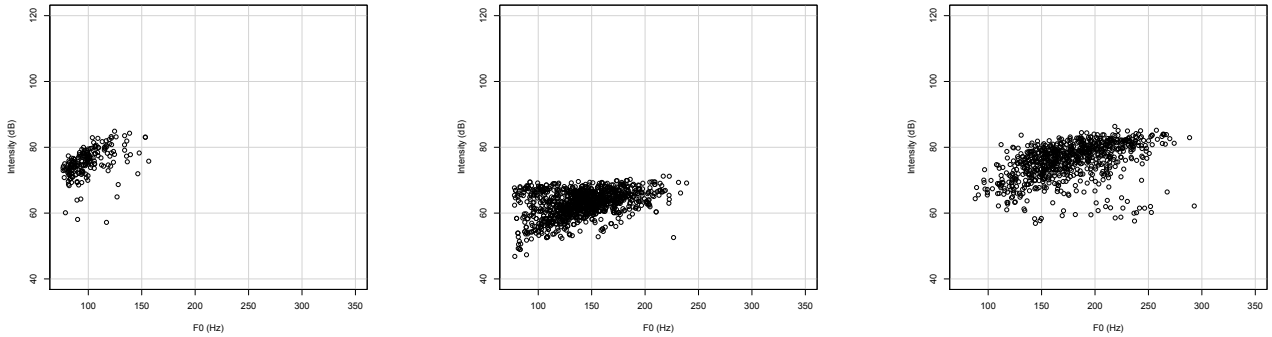

Figure 3: Speech Range Profiles of Brazilian speaker Luiz Inácio Lula da Silva (S-BP). Communication context: (a) informal interview addressed to the interviewer (INT); (b) formal conference addressed to other politicians (CON); (c) formal monologue addressed to followers (MON). X-axis: F0 (Hz), Y-axis: Intensity (dB).

6. References

[1] M. Weber, *The theory of social and economic organization.* New York, USA: Oxford University Press, 1920.

[2] S. A. Haslam, S. D. Reicher, and M. J. Platow, *The new psychology of leadership*. Hove, UK and New York, US: Psychology Press, 2011.

[3] R. Signorello, F. D'Errico, I. Poggi, and D. Demolin, "How charisma is perceived from speech. A multidimensional approach," in *ASE/IEEE International Conference on Social Computing* (IEEE Computer Society, ed.), (Amsterdam, The Netherlands), pp. 435–440, September 3-6, 2012 2012.

[4] Signorello, R., D'Errico, F., Poggi, I., Demolin, D., and Mairano, P. (2012). Charisma perception in political speech: a case study. In Mello, H., Pettorino, M., and Raso, T., editors, *Proceedings of the VIIth GSCP International Conference : Speech and Corpora*, pages 343–348. Firenze University Press.

[5] A. Lamarche, *Putting the Singing Voice on the Map*. PhD thesis, KTH School of Computer Science and Communication, Stockholm, Sweden, 2009.

[6] J. Catford, *A Practical Introduction to Phonetics*. Oxford, USA: Oxford University Press, 2002.

[7] Aristotle, *Rhetoric*. Retrieved from Amazon.fr: Acheron Press, Kindle version ed., trans. 1991.

[8] R Development Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012. ISBN 3-900051-07-0.

[9] R. Baken and R. Orlikoff, *Clinical Measurement of Speech and Voice*. San Diego, USA: Singular Publishing Group Inc., 2nd revised ed., 2000.

[10] I. Poggi, "The goals of persuasion," *Pragmatics & Cognition*, vol. 13, no. 2, pp. 297–336, 2005.

[11] Field, A., Miles, J., and Field, Z. (2012). *Discovering Statistics Using R*. Sage Publications Inc., London, UK.

[12] M. Zuckerman and R. Driver, "What sounds beautiful is good: The vocal attractiveness stereotype," vol. 13, no. 2, pp. 67–82, 1989.

[13] K. R. Scherer and U. Scherer, *Speech behavior and personality*, ch. 6, pp. 115–135. New York, USA: Grune & Stratton, j. darby ed., 1981.

[14] A. Rosenberg and J. Hirschberg, "Charisma perception from text and speech," *Speech Communication*, vol. 51, no. 7, pp. 640 – 655, 2009. Research Challenges in Speech Technology: A Special Issue in Honour of Rolf Carlson and Björn Granström.

[15] J. Ohala, "An ethological perspective on common cross-language utilization of F0 of voice," *Phonetica*, vol. 41, no. 1, pp. 1–16, 1984.