

# Modelling politicians' phono-styles with TTS

Carmen Patricia Pérez

CLILLAC-ARP, Université Paris Diderot – Paris 7, France

<https://doi.org/10.36505/ExLing-2017/08/0022/000324>

## Abstract

An oratory technique often observed in politicians consists in playing with prosodic features in order to build a particular phono-style. In a previous perceptual and acoustic analysis, I showed that Hugo Chávez (former president of Venezuelan) and José Zapatero (former prime minister of Spanish) have completely different phono-styles, which are easily recognizable by listeners. In this study, text-to-speech (TTS) synthesis was initially used to produce neutral utterances. The prosodic parameters of these utterances were then varied to reproduce Chávez and Zapatero's specific phono-styles. The reproductions were tested in a series of perceptual experiments which show that the modifications made to intonation patterns are enough to identify the two politicians.

Key words: intonation, phono-style, public speech, politician speech, synthesis.

## Introduction

A previous perceptive and acoustic analysis of Spanish-speaking politicians (Pérez 2016) revealed that H. Chávez (former president of Venezuela) and J. Zapatero (former prime minister of Spain) have completely different phono-styles in public 'spontaneous' speeches. In this study, I will first briefly describe Chávez and Zapatero phono-styles, then propose three perceptive tests using text-to-speech synthesis to convert written text into speech. F0, duration and chunking were modified to reproduce each politicians' prosodic patterns. The results of these tests show that Chávez and Zapatero are recognizable thanks to these stylized intonation pattern confirming the previous prosodic analyses.

## Chávez and Zapatero prosodic pattern

Chávez and Zapatero's phono-styles were established thanks to the acoustic analysis of 22 utterances, extracted from public 'spontaneous' speeches. It was observed that each politician cuts his speech into chunks (1982ms for Chávez and 1127ms for Zapatero on average) and makes large pauses between chunks (811ms for Chávez and 720ms for Zapatero on average). Chávez's phono-style consists of a repetitive contour at the end of each intonation phrase (IP). This contour rises on the stressed syllable (the penultimate one for 80% of Spanish words), followed by a high dive of about 16 semi-tones on the last unstressed syllable of the word. Furthermore, we find lengthening on this final unstressed syllable (two, three or four times the length of the preceding stressed syllable). Zapatero's

---

ExLing 2017: Proceedings of 8th Tutorial and Research Workshop on Experimental Linguistics, 19-22 June, Heraklion, Crete, Greece

prosodic pattern is characterized by small chunks and the use of a rising contour on the stressed syllable, continuing with a rise to a higher F0 value or floating tone on the last unstressed syllable. The politicians' schematized prosodic patterns are shown in fig.1 and 2.

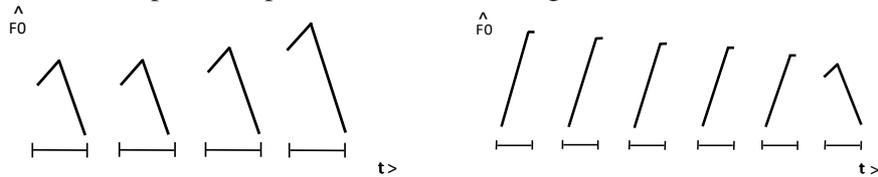


Figure 1. Chávez's prosodic pattern. Figure 2. Zapatero's prosodic pattern.

### Perceptual experiments: Methodology

**Materials and design:** The corpus consisted of four sentences adapted from the two politician's actual sentences, which were modified to be politically and dialectally neutral. They were produced using the TTS synthesis engine embedded in WinPitch. The utterances are 'prosodically neutral' (with a feminine voice) and were resynthesized according to the two politicians' respective prosodic characteristics. To ensure the prosodically modified speech was satisfactory, PSOLA generated pitch markers were carefully checked and edited.

In the first test, utterances were re-synthesized by modifying F0 parameter according to each of the above prosodic patterns. In the second test, F0 and duration of the last syllable of each chunk/IP were modified (the last unstressed syllable was twice as long as the stressed one). Each test had 12 synthesized utterances: 4 Chávez (C), 4 Zapatero (Z) and 4 'normal' (N) produced by TTS. Finally, in the third test, Chávez's utterances from the second test and Zapatero's utterances from the first were used. Four utterances were also added with shorter chunks for each politician. This test thus had 20 utterances: 4 C, 4 Z, 4 N, 4 C with shorter chunks (CSC) and 4 Z with shorter chunks (ZSC). Speed and pauses of the tests were not modified.

**Participants:** 10 native speakers of Spanish (Spanish or Venezuelan) and 10 non-Spanish-speaking participants (Chinese or Japanese mother tongue) who would, in a certain way, give a psychophysical (and not linguistic) perception.

**Procedure:** Participants started with a learning task of five minutes where they first listened to the unmodified Chávez and Zapatero recordings in order to get accustomed to their phono-style, followed by synthesized utterances in order to get used to synthetic female voice. Finally, they listened to re-synthesized Chávez and Zapatero utterances in order to test the phono-styles' recognition. Each test was based on a seven-step answer scale: surely C, probably C, maybe C, Other, maybe Z, probably Z and surely Z. Each step corresponded to a number: -3, -2, -1, 0, 1, 2, 3 respectively.

## Results

Test 1 shows that Chávez is not well recognized; Normal (Other) and Zapatero may be recognized, but most of the time confounded: the recognition variability is considerable (cf. SD on fig.1). I conclude that the F0 pattern alone is not enough to enable the speakers' differentiation.

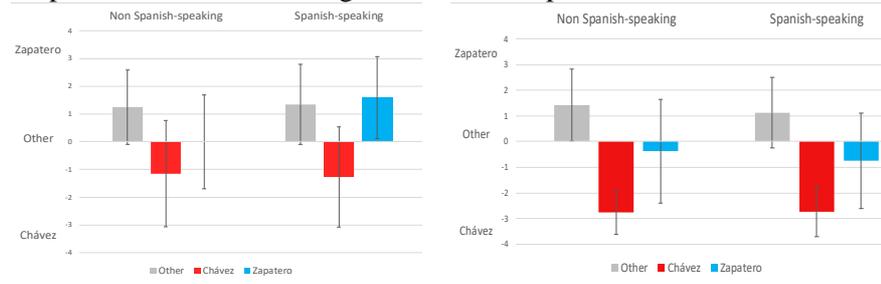


Figure 3. Test 1 (left) and 2 (right) results. Mean and SD of politicians' recognition.

Test 2, where the utterances had a lengthening of the IP last syllable, shows that Chávez is very well recognized; he is never confused with Zapatero or with normal utterances (cf. fig. 4). The normal condition is confused with Zapatero, while Zapatero is often confused with Chávez. I conclude that lengthening is a characteristic of Chávez and not of Zapatero.

Test 3 shows that the Chávez pattern is still very well recognized despite the shorter chunking. On the contrary, the shorter chunks help the Spanish native speakers to recognize Zapatero. Normal utterances are confused with Zapatero's but the mean is the smallest one of the three tests.

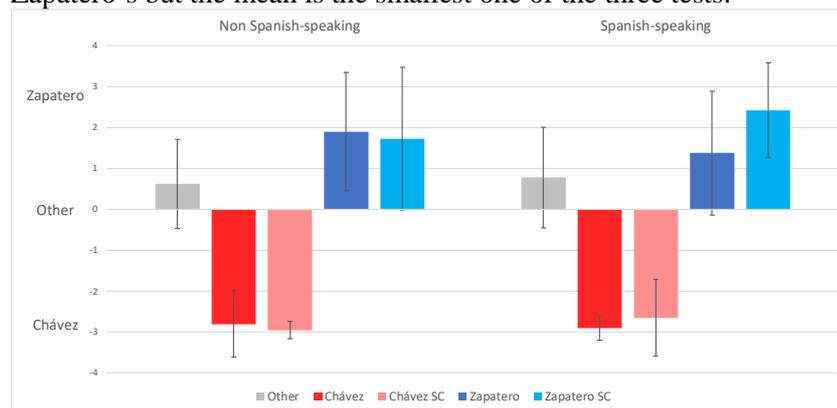


Figure 5. Test 3 results. Mean and SD of politicians' recognition.

## Conclusion

These synthesis experiments show that in general F0 modification is not sufficient to identify Chávez, who is well recognized only when the last IP syllable duration has been raised (test 2 and 3). Zapatero's pattern is often confused with the 'normal' one, probably because his prosodic pattern is similar to the 'standard' phono-syntactic Spanish intonation. Shortening the chunks provides very good Zapatero recognition by Spanish-speaking hearers. Results illustrates that I) the F0 movement and the duration of the last unstressed syllable is crucial in these politicians' identification and for oral Spanish and II) Chávez and Zapatero phono-style are easily recognizable regardless participants mother tongue.

## References

- Fónagy, I., 1983. *La vive voix. Essais de psycho-phonétique*. Paris: Payot.
- Léon, P., 1993. *Précis de phonostylistique. Parole et expressivité*: Nathan.
- Martin, P., 2010. Intonation in Political Speech: Ségolène Royal vs. Nicolas Sarkozy. Rome, pp. 54-64.
- Martin, P., 2015. *The structure of spoken language. Intonation in Romance*: Cambridge University Press.
- Pérez, C.P., 2016. A study of the phono-styles used by two different Spanish-speaking political leaders. *Proc. of Speech Prosody 8*, Boston, pp. 410-414.
- Sosa, J. M., 1999. *La entonación del Español: Su estructura fónica, variabilidad y dialectología*. Madrid: Catedra.
- Touati, P., 1995. Pitch range and register in French political speech. *Proc. XIII International Congress of Phonetic Sciences*, Volume 4, pp. 244-248.