PROSODIC MANIFESTATIONS OF POLITENESS IN PORTEÑO SPANISH
WH- INTERROGATIVES: TERMINAL CONTOURS, F0 MITIGATION AND SYLLABLE DURATIONS

Raoul GUCEK a & David LE GAC a, b

a DyLiS, Université de Rouen, b CNRS LLING, Université de Nantes
raougucek@gmail.com; david.le.gac@gmail.com

ABSTRACT

Politeness is often described as being prosodically cued through higher F0, as per the Frequency Code premises [12–14, 21, 25], and a slower speech rate [10, 11, 15, 16, 18, 20, 23, 27]. In Porteño Spanish, opposite findings [2, 6, 8] regarding wh-interrogatives’ pitch contours were attributed to a mark of politeness.

The analysis of 280 wh-interrogatives produced by nine speakers of Porteño Spanish, within colloquial and polite contexts, allowed us to determine that politeness does not actually favour higher F0. On the contrary, politeness is expressed by F0 mitigation through lowered overall pitch and reduced span. Our results also refute [26]’s hypothesis that interrogative adverbs would make utterances containing them more sensitive to politeness. In Porteño Spanish, adverbs do not pilot rising terminal contours to signal politeness. As for syllable durations, lengthening did not occur and speech rate kept constant in our corpus in the polite condition.

Keywords: wh-interrogatives, politeness, prosodic mitigation, terminal contour, Porteño Spanish

1. INTRODUCTION

In Spanish, the intonation of wh-interrogatives has been described as roughly corresponding to that of declaratives [1, 6, 22, 24]. More specifically, [2] and [8] explain that wh-words typically start the utterance carrying a high or rising accent (e.g. H* or L+H* in the ToBI system [7]), and the end of the interrogative is characterized by a progressive falling contour with the nuclear accent L* followed by an L% boundary tone. Yet, a rising final contour is reported by [19, 22] to soften the interrogation and make it sound polite. Recent studies on Porteño Spanish prosody [2, 6, 8] have obtained contradictory results regarding wh-interrogatives’ terminal contours. [2] and [8] have found that falling terminal contours were the most commonly used, whereas for [6] it is the rising ones. [8] suggested that H% might be predominant in oxytonic final words but also made various suggestions to explain these discrepancies: the lack of context in [6]’s study, different pragmatic goals (such as information seeking, inviting, criticizing, etc.) or the absence of control of sociolinguistic variables in all of the three studies. [6] also point out differences in the tasks proposed to the speakers: repetition of utterances in their study, reading out loud for [2] and semi-spontaneous speech for [8]. The latter is recognized by [15, 26] as possibly enhancing intonation and making it more discriminant among contexts.

In her study of the prosody of wh-interrogatives in Colombian Spanish, [26] also considers rising terminal contours as polite and further states that the grammatical category of the wh-word could be of influence. Specifically, adverbs seem to be capable of making the interrogatives containing them more sensitive to colloquial than polite contexts.

As for politeness, the separation line with formal speech is blurry and both terms are frequently employed as synonyms. [17] connects these notions, stating that the use of formal forms creates a formal atmosphere where participants, being kept away from each other, avoid imposition. Thus, to create this formal atmosphere is to be polite. This conjugated effort of preservation correlates to the facework [4, 9] that participants put into place in face-threatening acts. A way to achieve this – apart from the use of adapted vocabulary and syntax, for example – is through prosodic modulations, whose origins could be biological. The Frequency Code [12, 21] believes that the prosodic colloquial vs polite forms are based on a morphological-acoustic correlate: the bigger the larynx, the lower the pitch. This leads to a universal biological code, which associates high frequencies with submission and low frequencies with domination. [12] then extrapolates from this a paralinguistic meaning with an affective interpretation of these parameters: namely, friendliness and politeness are associated with ‘submissiveness’ and a higher pitch.

[15] compiled studies in various languages in favour of the Frequency Code (like [13] in Catalan, [14] in Peninsular Spanish and [25] in Japanese). However, they also list many others against it. For instance, [10] found that despite individual differences among speakers, average F0 measures...
seem to be lower in the formal register condition in German. In Korean as well, low pitch rather than high pitch was used to mark a polite register when addressing a person of superior status or age [27]. [15] and [16]'s studies in Catalan also show F0 mitigation (i.e. the lowering of the overall pitch) happening in polite speech.

As far as the other prosodic cues are concerned, it is reported that politeness is expressed in particular by a slower speech rate. This phenomenon has been consistently described [10, 11, 15, 16, 18, 20, 23, 27] with no contradictory result so far.

In this paper, we present the results of a pilot study on the prosodic expressions of politeness in wh-interrogatives in Porteño Spanish. In particular, we aim to determine whether, in accordance with the Frequency Code, the polite context favours the rising pitch contours at the end of wh-interrogatives as suggested [6, 19, 22, 26] or whether the wh-word category [26] and the stress pattern of the final word [8] play a role in the selection of the terminal contour. In addition to these local properties, we also investigate overall F0 and temporal characteristics of the wh-interrogatives such as the average pitch, the pitch range and syllable duration.

2. METHODOLOGY

2.1. Participants

Nine speakers of Porteño Spanish (5 females, 4 males; mean age 42.1, minimum 20, maximum 58) were recorded. They were all educated, either through academic studies or professional training.

2.2. Materials

The materials consisted in a corpus of 32 wh-interrogatives equally divided between a colloquial and a polite context, with each of these employing the same number of adverbs and pronouns. According to the Diccionario de la Real Academia Española [5], interrogative adverbs are cómo (how), cuándo (when), dónde (where) and por qué (why), and interrogative pronouns are cuál (which), cuánto (how much/long), qué (what) and quién (who). As for the stress pattern of the final words, the balance of (non-)oxytonic final words was roughly the same for each wh-word and formality condition.

The utterances were contextualized in order to carry a colloquial or polite value. The examples in (1) and (2) show discourse contexts encoding a colloquial situation and its polite correspondent, respectively, to elicit the target sentence (TS) ¿Cuál me queda mejor? (Which one suits me better?).

(1) Te vas de fiesta y no sabes qué chaqueta poner te, dudas entre dos. Se las enseñas a un amigo y le preguntas: TS  
(You’re going to a party and you’re thinking about which jacket to put on, you’re wavering between two. You show both of them to a friend and ask him: TS)

(2) Está en una tienda probándose trajes/vestidos para un casamiento. Le pregunta al vendedor: TS  
(You are in a shop trying on a dress/suit for a wedding. You ask the shop assistant: TS)

Two of the three pragmatic factors of [15] were implemented in the discourse contexts presented to the participants: the social distance between the interlocutors and their relative power. The third one, the cost of the action, does not present any opportunity to be tested using wh-interrogatives.

2.3. Experimental procedure

The speakers were recorded in Buenos Aires, Argentina. The recordings took place in quiet rooms with a Zoom H1n microphone, in stereo at 24-bit/48 kHz and with a saturation limiter activated.

Each speaker was presented with plasticized cards in different randomized orders, containing a TS introduced by the description of the context. To avoid boredom and fatigue, only half of the corpus was offered to each of them. This represented one of four TSs introduced by an adverb and one of four of those introduced by a pronoun, induced in the colloquial context – which was always offered first – and then their eight polite correspondents.

Participants were invited to ask any question that would be necessary to clarify the contextualization. The readings were systematically repeated once in order to discard possible faulty ones. In the end, 280 utterances were exploitable.

2.4. Labelling, measurements and quantitative analysis

The TSs were segmented and labelled using Praat software [3]. They were manually segmented into syllables and an F0 measurement was taken at times ⅓, ⅔ and ⅔ of the duration of each syllable. From these measurements, the following data were established for each entire wh-interrogative/TS:
1. average pitch, min and max F0 values in Hz;
2. F0 value on the middle of the stressed syllable of the wh-words;
3. F0 value at the end of the TSs, which represents the F0 target of the final boundary tone of the sentence;
4. pitch range in semitones (ST) defined as the difference between the max and min F0 of the TSs (the min F0 was taken as the reference value);
5. syllable duration of the wh- words and of the last two syllables of the TSs;
6. speaking rate over the entire TSs defined as the duration of the interrogation divided by the number of syllables.

As for the pitch contours, they were annotated in an independent tier following [7]’s Sp_ToBI annotation system, as shown in Fig. 1:

Figure 1: Illustration of annotations for the TS ¿Dónde dejo el vino? (Where should I leave the wine?), in colloquial context and introduced by an adverb.

![Pitch range illustration](image)

3. RESULTS

3.1. Nuclear pitch configurations

Each nuclear pitch configuration (NPC), i.e. the nuclear pitch accent plus boundary tone(s) for each wh- interrogative, was annotated. Table 1 below presents the distribution of the NPCs in the two formality conditions. It also includes the grammatical category of the wh- word to establish this distribution.

Table 1: Inventory of NPCs within colloquial and formal contexts in % and actual number; “Adverbs” and “Pronouns” indicate the category of the wh- word.

<table>
<thead>
<tr>
<th>NPC</th>
<th>COLLOQUIAL</th>
<th>POLITE</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
</tbody>
</table>

The distribution of the NPCs is similar in colloquial and polite contexts. More specifically, L* L% is the most widely used contour in both formality conditions (63.9% in the colloquial condition and 72.1% in the polite one). Slightly more L* L% occurred in the polite context but this difference cannot be considered significant. Overall, the falling NPCs (i.e. those consisting in a pitch drop at the end of the sentence) clearly outnumber the rising ones, which end with a slight (L* M%) or a steep F0 rise ((H)H%): there is 73.6% of falling NPCs vs 26.4% of rising NPCs in the colloquial condition and 79.4% vs 20.6% of each contour type in the polite one. The results thus indicate that the politeness condition does not favour a terminal rising pitch.

The grammatical category of the wh- word does not have any influence either over NPCs. Table 1 shows that interrogative adverbs and pronouns generate equivalent quantities of rising contours in both formality conditions.

With regard to [8]’s (p. 290) claim that H% boundary tone typically occurs in questions ending in oxytonic words, the overall results do not seem to support this finding. As can be seen in Fig. 2, the distribution of the boundary tones is quite similar at the end of the oxytonic and non-oxytonic words. Nevertheless, if one puts aside the predominance of L% and considers the cases where speakers have not chosen the L% tone, there seems to be a preference for the (H)H% categories at the end of the oxytonic words (17.6%/22.9%=76.9%), while the M% seems to be preferred with non-oxytonic words (15.7%/23.6%=66.5%). However, it is not yet possible to conclude because of insufficient data.

Figure 2: Distribution of the boundary tones depending on the (non-)oxytonicity of the final word.

![Boundary tone distribution](image)

3.2. Overall pitch characteristics of wh- interrogatives

Table 2 below presents the mean F0 and standard deviation (SD) for the six pitch values defined in 2.4. in the two formality conditions. The table also contains the results of a series of mean comparisons.
(t-tests). For these measurements and mean comparisons, we have selected only TSs that were strictly identical (indeed, some TSs had to be formulated differently according to the context to ensure coherence and authenticity). This amounted to 134 utterances in each context.

Table 2: Mean F0 and standard deviation (SD) for the 6 pitch values defined in 2.4. in the two formality conditions. “Inter.” and “Fin.” refer to the F0 value on the wh-word and to the final boundary tone of the TSs, respectively. The table also includes the results of a series of t-tests (the degrees of freedom are 133 in all cases)

<table>
<thead>
<tr>
<th>Pitch Range</th>
<th>Colloquial</th>
<th>Polite</th>
<th>T-test results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average pitch (Hz)</td>
<td>199.00 72.18</td>
<td>187.84 63.70</td>
<td>t = 5.822 p &lt; .001</td>
</tr>
<tr>
<td>Min. (Hz)</td>
<td>142.24 48.75</td>
<td>136.30 43.75</td>
<td>t = 3.416 p &lt; .001</td>
</tr>
<tr>
<td>Max. (Hz)</td>
<td>271.17 106.52</td>
<td>251.05 95.22</td>
<td>t = 5.346 p &lt; .001</td>
</tr>
<tr>
<td>Inter. (Hz)</td>
<td>244.66 92.56</td>
<td>226.36 83.12</td>
<td>t = 5.216 p &lt; .001</td>
</tr>
<tr>
<td>Fin. (Hz)</td>
<td>162.16 74.51</td>
<td>149.44 57.46</td>
<td>t = 2.944 p = 0.004</td>
</tr>
<tr>
<td>Pitch Range (ST)</td>
<td>10.79 3.24</td>
<td>10.26 3.59</td>
<td>t = 1.795 p = 0.075</td>
</tr>
</tbody>
</table>

As can be seen, in the polite condition, all F0 values (average pitch, min, max, Inter. and Fin.) are significantly lower (p<0.001, p=0.004 for Fin.) than those in the colloquial condition. This suggests that F0 mitigation in the polite context affects the whole wh-interrogative and not a specific contour or word. The only non-significant difference between contexts is the pitch range, which corroborates the idea that it is the overall F0 pattern of the sentences that is lowered in the polite context.

As regards the location of the F0 maxima, [24] (p. 216) pointed out that the highest peak tends to coincide with the wh-word (interrogative pronoun or adverb), although not necessarily with its accented syllable, without dialectal distinctions. In Porteño Spanish, the results show that the F0 max is not located on the stressed syllable of the wh-word in most cases. This may well be a case of late tonal alignment (as in Fig. 1 with L+>H*), which will require further investigation.

3.3. Syllable durations

Table 3 presents the mean duration of the last two syllables of the wh-word and of the TSs in the two formality conditions. It also gives the speaking rate of the TSs (i.e. duration of the TS divided by the number of syllables, cf. 2.4.).

Unlike the findings of [10, 11, 15, 16, 18, 20, 23, 27], Table 3 shows that politeness does not affect either syllable duration or speech rate in Porteño Spanish. No significant differences can be seen for the mean durations between the two formality conditions.

Table 3: Mean duration in seconds of the wh-word (“Inter.”), the last two syllables of the TSs (“Fin.”) and the syllables over all the TSs (representing the speaking rate) in the two formality conditions. The table also includes the results of a series of t-tests

<table>
<thead>
<tr>
<th></th>
<th>Colloquial</th>
<th>Polite</th>
<th>T-test results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inter. (s)</td>
<td>0.253 0.101</td>
<td>0.262 0.112</td>
<td>t = -1.578 p = 0.117</td>
</tr>
<tr>
<td>Fin. (s)</td>
<td>0.472 0.125</td>
<td>0.465 0.121</td>
<td>t = 0.884 p = 0.379</td>
</tr>
<tr>
<td>Syllables (s)</td>
<td>0.174 0.04</td>
<td>0.176 0.041</td>
<td>t = 0.758 p = 0.45</td>
</tr>
</tbody>
</table>

4. DISCUSSION AND CONCLUSION

In this paper we aimed to determine what the prosodic manifestations of politeness are in wh-interrogatives in Porteño Spanish. Our results show that politeness has no influence on NPCs. Indeed, a final L% occurs in the majority of the interrogatives (more than 75%) no matter the formality condition. This means that politeness does not favour high or rising NPCs, contrary to [6, 19, 22]’s proposals. The grammatical category of the wh-word has no influence either on the NPCs, in contrast to [26]’s allegation in Colombian Spanish; neither does the stress pattern of the final word as claimed by [8].

As for global F0 values (average pitch, min and max F0, Inter. and Fin.), mitigation occurs in all of these parameters (except for the pitch range) in the polite condition. Our results thus contradict previous research and in particular the Frequency Code premises, which predict that high or rising F0 is a cue for politeness. They are in line with the studies of [15, 16] for Catalan, [10] for German and [27] for Korean, who also found that F0 mitigation occurred in the polite condition.

Another robust prosodic cue for politeness reported in earlier studies [10, 11, 15, 16, 18, 20, 23, 27] is a slower speech rate. Unexpectedly, we did not find this tendency in our data. The duration of the wh-words and that of the last syllables of the sentences exhibit no significant difference between the formality conditions. Likewise the speech rate remains the same in both contexts.

Further research on wh-interrogatives’ prosody in Porteño Spanish still needs to work out how NPCs are chosen between rising and falling. The location of the F0 max is of interest too, as it appears that the F0 max is not aligned with the stressed syllable of the wh-word in Porteño Spanish. Finally, and as proposed by [15, 26], the present results would benefit from being tested with a different elicitation task, such as analysing (semi-) spontaneous speech.

5. ACKNOWLEDGEMENTS

We thank the participants of the study and David Imbert (CNRS LLING, Université de Nantes) for his help and advice with the statistical tests; all remaining errors are ours.
6. REFERENCES


