The multiple prosodic cues differentiating questions and statements in Miami Cuban Spanish

Ann M. Aly

University of California, Los Angeles
ann.m.aly@ucla.edu

ABSTRACT

Previous work on the intonational phonology of Miami Cuban (MC) Spanish [1, 2, 3] has revealed that broad focus declaratives and absolute interrogatives have the same nuclear tones. In laboratory speech, [2] found that these two sentence types, when compared paradigmatically, show significant differences in peak scaling and pitch range [2]. The present study examined peak scaling with syntagmatic comparisons between nuclear and prenuclear peaks as well as paradigmatic comparisons of pitch range and f0 peak timing in laboratory and semi-spontaneous speech. The results reveal significant differences in peak scaling and pitch range, but no differences in peak timing, confirming [2]. However, preliminary perception results reveal no difference in listener interpretation when only peak scaling is manipulated, suggesting free variation or additional cues to sentence modality in this dialect.

Keywords: Miami Cuban Spanish, Intonational Phonology, Dialects, Autosegmental-Metrical

1. INTRODUCTION

The Autosegmental-Metrical (AM) model of intonational phonology of Miami Cuban (MC) Spanish established in [3] revealed a relatively small tonal inventory of contrastive pitch accents and boundary tones when compared with other dialects of Spanish, such as those described in [14]. Whereas the AM models of Pan Spanish [8] and other Caribbean dialects of Spanish, such as Puerto Rican [4] and Dominican [16] contain at least six pitch accents and four boundary tones (including monotonal and bitonal boundary tones), MC Spanish contains only four pitch accents and three monotonal boundary tones. Table 1 gives the phonetic description of the tonal inventory in MC Spanish proposed in [3].

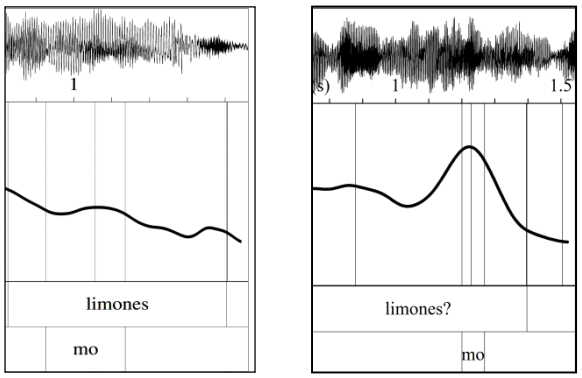
Table 1: Phonetic descriptions MC Spanish tonal inventory

Pitch accents	Phonetic description
L*+H	Rising with f0 trough in tonic syllable
L+H*	Rising with f0 peak in tonic syllable
H*	f0 peak in tonic syllable
L*	f0 peak in tonic syllable
Boundary tones	Phonetic description
L%	Falling f0 after last tonic syllable
M%	Slight f0 fall after last tonic syllable
H%	Rising f0 after last tonic syllable

Because the majority of Spanish dialects described in [14, 8] have predictable or default prenuclear pitch accents, sentential meaning determined by prosody (and not by syntax or morphology) is relegated to the nuclear contour of the utterance, which consists of a combination of the last (i.e., nuclear) pitch accent (NPA) and boundary tone in the Intonational Phrase (IP). Due to the overall small tonal inventory of MC Spanish, various sentence types are realized with the same nuclear contour. For example, a rise-fall nuclear contour typically represents broad focus statements, absolute interrogatives [1, 2, 3] as well as imperatives, exclamatives and narrow focus statements [3], although falling declaratives (H+L* L%) were also reported in the latter study.

Although sentence types such as wh-questions and imperatives are typically disambiguated lexically in Spanish with an interrogative word or imperative verbal inflection, broad focus declaratives are often syntactically and lexically identical to absolute interrogatives, being differentiated only by intonation. This suggests that additional prosodic or acoustic cues must be present to help speakers disambiguate sentence types with similar intonation, such as the scaling or timing of the NPA peak, as reported for Neapolitan Italian in [7] or higher pitch range for questions in Puerto Rican Spanish [4] and Hungarian [11]. Indeed, [2] found that questions in MC Spanish have a higher pitch range then declaratives. An example is shown in Figure 1 in which the nuclear contour of a declarative (left) has a narrower pitch range than that of a question (right). These are from the utterance *Lorena mide limones* (‘Lorena measures lemons/limes’).

Figure 1: Nuclear contours of a declarative (left) and a question (right).



Based on laboratory speech of MC Spanish speakers, [1, 2] proposed an L*+H prenuclear pitch accent and L+H* L% nuclear contour for both broad focus statements and declaratives, but showed that these contours are realized differently. When measured, prenuclear and nuclear f0 peaks and valleys (from which pitch range was also inferred) as well as the timing (or alignment) between syllable onset and the start of the pitch accent rise, a significant difference was found for pitch scaling only, in which questions had higher pitch peaks in both prenuclear and nuclear pitch accents when compared to those in declaratives.

The present study seeks to further investigate the acoustic cues that may play a role in disambiguating broad focus statements (henceforth 'declaratives' and absolute interrogatives (henceforth 'questions') in MC Spanish, by examining several factors that were not considered before. First, whereas [2] showed consistent and significant effects in laboratory speech, the present study considers semi-spontaneous speech, as differences in intonation due to task type have been reported for Spanish declaratives in [9]. Second, the syntagmatic comparison between declarative and question peaks performed in [2] would not capture the relationship between prenuclear and nuclear tonal targets in the same sentence. Therefore, the present study will make paradigmatic comparisons between tonal targets within the same sentence type as well as the syntagmatic comparisons. Finally, the present study reports preliminary results of a perception experiment designed to test how native speakers interpret these cues.

2. EXPERIMENT 1: PRODUCTION

2.1. Participants

Nine speakers of MC Spanish (four male, five female) participated in the present study. Four speakers were born in Cuba and currently reside in southern Florida (Miami-Dade or Lee County) and five speakers were born in Miami, Florida. Cuba-born participants had resided in the United States for an average of 25 years and were a median age of 52.2 years; Miami-born participants were second-generation Cuban-Americans whose median age was 30 years. All participants reported using both Spanish and English on a daily basis.

2.2. Procedure

Participants completed three different tasks: a reading task, designed to elicit more careful speech and control for number of syllables, stress placement, and sentence type (such as imperative or declarative); a Discourse Completion Task (DCT) adapted for MC Spanish from [14], which elicits numerous sentence categories with prompts that allow participants to answer freely; and a structured interview about the participant's daily life and opinions on current events, which elicited more naturalistic speech that still spanned several sentence types. All participants were recorded in a quiet location with the researcher with an Olympus LS-11 portable voice recorder (16 bit rate, 44.1 kHz).

The data collected was analysed in *Praat*, version 5.3.73 [6]. Prosodic events were transcribed using the Spanish Tone and Break Indices (Sp_ToBI) conventions established for MC Spanish [3]. In addition, to quantify the peak alignment and scaling, the following measurement were made: prenuclear and nuclear pitch accent peak height (in Hz and semitones), duration between the f0 peak to the offset of the associated (tonic) syllable of the NPA (in milliseconds), and pitch range (in Hz and semitones) of syntactically and lexically identical declaratives (n = 26) and questions (n = 28) with nuclear contours of L+H* L%.

A generalized linear mixed-effects model was implemented in R using the lme4 package [5, 13] to predict the probability of the sentence modality (question or declarative) outcome given the scaled independent variables peak timing, peak scaling, and pitch range with random intercepts for speaker.

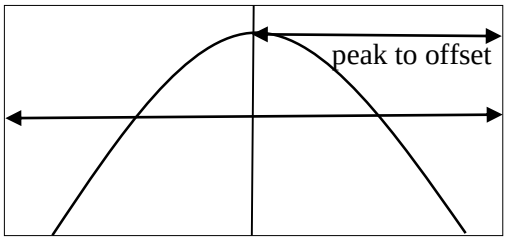
2.3. Results

2.3.1. Peak timing

Peak timing was measured in declaratives and questions as the duration in milliseconds (ms) between the highest f0 of the NPA (peak) and the end of the associated syllable. Figure 2 shows an example of these measurements, with arrows representing the locations of the duration measurements taken. In the event that there was no

clear pitch peak, but more of a plateau, the middle of the plateau was selected as the peak.

Figure 2: Schematic of peak timing measurements

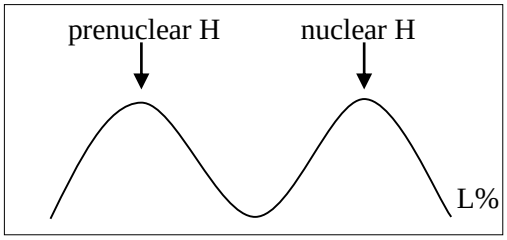


Due to the natural variation in speech rate between participants as well as syllable length and structure, the results of peak timing will be reported in both ms and as a proportion of the distance between the pitch peak and the end of the syllable divided by the duration of the entire syllable (2nd arrow in Fig.2). The peak timing results revealed that declarative pitch peaks occur earlier in the syllable (mean 121 ms before end of syllable or 57% of total syllable duration) than questions (mean 96 ms before end of syllable or 44% of total syllable duration). The statistical analysis revealed that this difference was not significant, Est = -.16, z = -.45, p = .65.

2.3.2. Peak scaling

Peak scaling was measured in declaratives and questions as the difference in f0 between the NPA peak and the immediately preceding prenuclear high target of the IP. As with pitch timing, results will be reported in both Hz and semitones (st) in an attempt to normalize the natural variation in pitch range between speakers. Figure 3 illustrates the locations at which f0 was measured for prenuclear and nuclear peak scaling.

Figure 3: Schematic of pitch scaling measurements

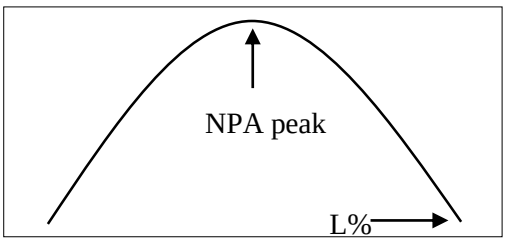


The results for peak scaling revealed that declaratives have a smaller f0 difference between NPA peak and the preceding prenuclear high target in comparison to questions, which showed a greater difference in f0 between these two tonal targets. The mean difference between prenuclear and nuclear high tonal targets was .73 st (8 Hz) for declaratives and 3.7 st (58 Hz) for questions, which was statistically significant, Est = 2.04, z = 2.66, p = .007.

2.3.3. Pitch range

The final variable, pitch range, was defined in the present study as the difference between the NPA peak and the low (L%) boundary tone minimum. As with the peak scaling measurements, pitch range is reported in Hz and st in order to normalize pitch range variation between speakers. Figure 4 shows the locations in the nuclear contour at which the f0 was measured.

Figure 4: Schematic of pitch range measurements



These measurements revealed that questions have a larger pitch range than declaratives, with questions containing an average difference of 11.8 st (151 Hz) between NPA and L% and declaratives containing an average difference of 9.6 st (82 Hz). This difference reached statistical significance, Est = 2.59, z = 3, p = .002.

To summarize, the results of the three variables discussed revealed the following: question NPA peaks occur later in the tonic syllable than those in declaratives, though this difference did not reach significance; the difference between NPA peak height and prenuclear peak height is significantly larger in questions than declaratives; and finally, the pitch range of questions is significantly larger than that of declaratives.

3. EXPERIMENT 2: PERCEPTION

3.1. Procedure

In order to test whether the perceptual weight of cues such as peak scaling or timing can be determined, a small-scale perception experiment was designed based on the data collected in Experiment 1. The measurements performed on the data collected in Experiment 1 revealed a smaller difference between prenuclear and nuclear pitch

peaks in declaratives (8 Hz, compared to 58 Hz in questions), which provided the basis for manipulating pitch scaling in Experiment 2.

Twelve questions served as "base utterances" that were manipulated using *Praat* so that the NPA peak was lowered (downstepped) to an f0 below the last prenuclear peak of the IP. The manipulated utterances were split in two versions of Experiment 2, resulting in six downstepped and six distractor utterances per version (also manipulated questions), which will serve as target utterances for future analysis. Participants were asked to listen to each utterance and choose from the multiple choice answers provided based on their first intuition. For example, the utterance *Marina bebe limonada* ("Marina drinks lemonade") would present the choice for declarative modality as "The speaker is telling someone else that Marina drinks lemonade" and the choice for question modality as "The speaker is asking someone if Marina drinks lemonade". Participants could also choose a "None of the above" option.

3.2. Participants

Both versions of the experiment were distributed as a Google Forms online survey to self-reported speakers of MC Spanish who were either born in or lived in Miami for at least ten years. Six participants (3 for each version) completed this pilot experiment.

3.3. Results

The results of the pilot experiment revealed that MC Spanish speakers still interpreted questions as questions even when the NPA peak is downstepped: Of 36 total responses, 25 responses categorized the altered (downstepped) utterances as questions, whereas only 8 categorized them as declaratives (3 responses were "None of the above"). Although a small sample, this pilot study suggests that a larger portion of the utterance be investigated for sentence modality, as will be discussed in Section 4.

4. DISCUSSION

The current study revealed several noteworthy findings. The comparisons between declaratives and questions in the present study involved semi-spontaneous data as well as both paradigmatic (pitch scaling) and syntagmatic (peak timing and pitch range) measurements, however, the effect for larger overall pitch excursions in questions corroborated the paradigmatic measurements performed in [2], a characteristic typical cross-linguistically in questions [15]. Additionally, the present study found no significant effect for peak timing between questions and declaratives, which was also found in the laboratory data in [2].

The results of the pilot study assessing the perceptual weight of the NPA peak scale revealed that manipulating the nuclear region of questions with downstepped peaks does not change listeners' interpretation of sentence modality, even in the presence of significant peak height differences between declaratives and questions. This suggests that the prenuclear region also contains cues to question modality, even if phonologically identical. Similar results have been reported for North Standard German [12], which reports that the prenuclear region indicates modality and the nuclear contour in this dialect reflects attitudinal information; and for Castilian Spanish [10], in which prenuclear peak height can define modality in the absence of nuclear region cues.

If the difference found in [2] and the present study for peak scaling is not recognized categorically by native speakers of MC Spanish, it is possible that both questions and declaratives are phonologically identical and the differences measured represent free variation in this dialect. However, if additional perception tests that consider the entire utterance reveal categorical differences in either the prenuclear or nuclear regions, then different phonological labels might be necessary to represent these differences (e.g., L+¡H* for upstepped NPA peaks in questions). Additional data from MC Spanish is needed in order to conclude the strength of these cues and the region (prenuclear and nuclear) to which they belong.

5. CONCLUSION AND FUTURE DIRECTIONS

The present study presented laboratory and semi-spontaneous data from nine speakers of Miami Cuban (MC) Spanish and investigated acoustic cues that differentiate questions and declaratives, which are not only syntactically and lexically identical, but phonologically identical with respect to intonation. Measurements of peak timing, scaling, and pitch range revealed results similar to those of [2] for MC Spanish; that is, that peak scaling and pitch range are significantly higher for questions than declaratives, whereas peak timing is not a significant predictor of sentence type, even with the inclusion of semi-spontaneous data. Preliminary perception results, however, suggest that additional research is needed to investigate the location of important cues to modality in an utterance, which is currently in process with a larger-scale perception experiment which takes the tonal events across the entire utterance into consideration.

6. REFERENCES

[1] Alvord, S. 2006. Spanish intonation in contact: The case of Miami Cuban bilinguals. University of Minnesota PhD Dissertation.

[2] Alvord, S. 2009. Disambiguating declarative and interrogative meaning with intonation in Miami Cuban Spanish. *Southwest Journal of Linguistics 28*(2): 21-66.

[3] Aly Bailey, A. 2014. Intonational Phonology of Miami Cuban Spanish: An AM Model. University of California, Los Angeles M.A. thesis.

[4] Armstrong 2010. Puerto Rican Spanish intonation. *In* P. Prieto and P. Roseano [Eds], *Transcription of Intonation of the Spanish Language*, 155-189, Lincom: Europa München.

[5] Bates, D., Maechler M., Bolker, B., Walker, S. 2014. lme4: Linear mixed- effects models using Eigen and S4. R package version 1.1-6.

[6] Boersma, P., Weenink, D. 2014. *Praat: Doing phonetics by computer* [computer program], version 5.3.73, retrieved from www.praat.org

[7] D'Imperio, M. 2002. Tonal alignment, scaling and slope in Italian question and statement tunes. *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence* (TIPA), 21, pp.25-44.

[8] Estebas Vilaplana, E, Prieto, P. 2008. La notación prosódica del español: Una revisión del Sp_ToBI. *Estudios de Fonética Experimental*, 17: 265-283.

[9] Face, T. 2003. Intonation in Spanish declaratives: Differences between lab and spontaneous speech. *Catalan Journal of Linguistics 2*: 115-131.

[10] Face, T. 2007. The role of intonational cues in the perception of declaratives and absolute interrogatives in Castilian Spanish. *Estudios de fonología experimental 16*, 185-225.

[11] Gósy, M., Terken, J. 1994, Question marking in Hungarian: timing and height of pitch peaks, *Journal of Phonetics 22*:269-281

[12] Petrone, C., Niebuhr, O. 2013. On the intonation of German intonation questions: The role of the prenuclear region. *Language and Speech 0(0): 1-39.*

[13] R Core Team. 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

[14] Roseano, P., Prieto, P. (Eds). 2010. *Transcription of Intonation of the Spanish Language*, Lincom Europa: München.

[15] Thorsen, N. 1983. Two issues in the prosody of Standard Danish. *In* A. Cutler and D. R. Ladd (eds.) *Prosody: Models and Measurements*. Heidelberg: Springer. pp. 27-38.

[16] Willis, E. 2010. Dominican Spanish Intonation, *In* P. Prieto and P. Roseano [Eds], *Transcription of Intonation of the Spanish Language*, 125-153, Lincom Europa: München.