



The intonation of echo *wh*-questions in Ecuadorian Spanish

Clara Huttenlauch¹, Sophie Egger¹, Daniela Wochner¹, Ingo Feldhausen²

¹ Department of Linguistics, University of Konstanz, Germany

² Institute for Romance Languages and Literature, Goethe University Frankfurt, Germany

{clara.huttenlauch|sophie.egger|daniela.wochner}@uni-konstanz.de,
feldhausen@em.uni-frankfurt.de

Abstract

In Romance languages, neutral echo *wh*-questions and counterexpectational (i.e., non-neutral or incredulous) echo *wh*-questions are usually distinguished by tonal differences in the nuclear contour. In this study, we show that Ecuadorian Spanish does not differentiate between those question types in terms of tonal targets, but by expanding the pitch range in the nuclear region. Results drawn from a production experiment based on semi-spontaneous speech with nine native speakers show that neutral as well as non-neutral echo *wh*-questions are realized with the same nuclear configurations, typically a rising one (L* H%). However, counterexpectational *wh*-questions are realized with a significantly larger pitch range than their neutral counterparts (with a difference of 2.03 semitones). From a diatopic point of view, our data indicate that there are prosodic differences in the realization of questions between the Andean and the coastal region. Speakers from the latter region realize rising as well as low and falling nuclear configurations (L* H%, L* L% and H* L%) and use a lower mean pitch. We hereby provide intonational support for the well-known division between *tierras bajas* ‘lowlands’ and *tierras altas* ‘highlands’ used for capturing the dialectal phonological variation on the segmental level of Latin-American Spanish.

Index Terms: echo *wh*-questions, incredulity, intonation, pitch range, Ecuadorian Spanish, dialectal variation, Sp_ToBI

1. Introduction

Echo questions are subclasses of questions that fully or partially repeat a prior utterance. Their main function is for the addressee to clarify the content of the preceding utterance and to elicit missing information when he or she failed to understand what the speaker just expressed in the discourse [1] (see [2] and [3] for additional functions). In a neutral echo *wh*-question (see (1) below), the hearer did not comprehend the name *Pedro* properly and asks who came. However, the question in (1) can also be interpreted as a non-neutral echo question. In this case, the hearer did understand the information properly, but is surprised about what has been said since it contradicts his or her expectations (namely, that *Pedro* was the one who came). Hence, he or she checks who came.

Echo *wh*-questions (1)
Context: *Pedro vino.* (‘Pedro came’)
Echo: *¿Quién vino?* (‘Who came?’)

As (1) illustrates, the two types of echo *wh*-questions are not distinguished syntactically, i.e., by alternations in word order (see [2] and [4] for details on the possibility of using *wh*-

in-situ and the quotative particles *que* ‘that’ and *si* ‘whether’). In many Spanish varieties the difference between these two types of questions is marked in terms of intonation. In Mexican Spanish, for example, neutral echo *wh*-questions (NEQs) have a rising nuclear configuration (L* LH%), while non-neutral or counterexpectational echo *wh*-questions (CQs) are characterized by a rising-falling contour (L+H* L%; [5]). In Venezuelan Spanish, NEQs have a falling nuclear contour (H+L* L%) and CQs are realized by a rising contour (L+H* H%; [6]). In Argentinian Spanish, NEQs have a falling boundary tone, while CQs end at a mid-level (L+H* HL% vs. L+H* M%; [7]). In Castilian Spanish, a (rising-)falling contour is attested for NEQs ((L+)_iH* L%) and a rising contour for CQs (L+_iH* HH%; [8]). The Castilian patterns have already been described in the seminal work of Escandell-Vidal [2], [3] and [4], in which she offers detailed analyses of the interrelation between syntax, semantics, pragmatics and the intonation of echo-questions. Next to the tonal differences, all varieties also distinguish NEQs from CQs by expanding the pitch range of the CQs ([5], [6], [7] and [8]). The combination of tonal movement and pitch excursion is also attested in other Romance languages, such as Italian [9] or Catalan [10].

In Ecuadorian Spanish, however, no differences in the tonal realization have been attested so far [11]. Both question types are characterized by a rising-falling nuclear configuration (L* HL%). While this strongly contrasts with the pattern previously described for other Spanish and Romance varieties, Ecuadorian Spanish also makes use of pitch excursion [11]. CQs can be marked by an extended peak value of the high tone within the nuclear configuration. Despite the thorough study of [11], the overall concept of the edited book in which the work appeared entailed that [11] analyzed only a very limited dataset: Two subjects from Quito uttered up to two echo *wh*-questions per type. For this reason, we wonder whether the inexistence of tonal differences can be replicated by means of a more broadly based study or whether the findings are rather a result of the too restricted sample. Two hypotheses (H1 and H2) guide our study:

- H1: Both NEQs and CQs are realized with a low nuclear pitch accent (L*) and a rising-falling contour at the end of the utterance (HL%; in accordance with [11]).
- H2: In contrast to NEQs, CQs may have a wider pitch expansion within the nuclear configuration (again, in accordance with [11]).

2. Experiment

To analyze the intonation patterns of CQs and NEQs in Ecuadorian Spanish we conducted a production experiment based on semi-spontaneous speech gathered by means of a

Discourse Completion Task [12] and [13]. In the prosodic analysis, we concentrated on the nuclear configuration (i.e., the nuclear pitch accent and the sentence-final boundary tone), the pitch range and the mean pitch throughout the utterance.

2.1. Methods

2.1.1. Participants

We recorded nine monolingual native speakers of Ecuadorian Spanish (four male) aged between 19 and 38 years (average 25.9 years, SD = 7.0). All of them participated voluntarily for a small reimbursement and they were unaware of the purpose of the study. Five of them originated from Quito, the capital of Ecuador, and one from Tulcán. Both cities are in the Andean region, henceforth, the highlands. The last three speakers grew up at the coast in Guayaquil, henceforth, the lowlands. All participants had been living/studying in Germany at the time of the recordings (range between 4 months and 12 years) and are learning German as a foreign language.

2.1.2. Materials

In order to compare the nuclear configuration and the pitch range of CQs and NEQs we generated 12 pairs of short contexts, one evoking the production of a non-neutral question, the other one favoring a neutral reading (see (2) and (3) below). Contexts that elicit CQs were constructed in parallelism to the one for counterexpectational echo *wh*-questions used in the questionnaire of the *Atlas of Spanish Intonation* [14]. Comparable segmental environments are desired to allow for a thorough prosodic analysis. Thus, we created contexts to evoke neutral *wh*-questions with similar words and parallel syntax as in the non-neutral condition. Possible target utterances elicited by contexts (2) and (3) are shown in brackets. The contexts were controlled for eliciting mostly sonorant material within the region of the nuclear configuration (such as *(se) llama* ‘is called’ in the given examples) in order to allow for a better pitch tracking.

CQ: (2)

Te comentan que la madre de un conocido se llama Zoila Vaca. No puedes creer que alguien se llame así y preguntas muy incrédulo cómo se llama. (Possible answer: ¿Cómo se llama?)

‘You are told that the name of a friend’s mother is Ima Cow (pronounced: I’m a cow). You can’t believe that this is someone’s name and you ask incredulously what her name is.’ (Possible answer: What’s her name?)

NEQ: (3)

En una fiesta te presentan a la tía de tu novio/novia pero no sabes si has entendido bien el nombre. Pregúntale cómo se llama (Possible answer: ¿Cómo se llama?)

‘At a party, you are introduced to your partner’s aunt but you are not sure whether you understood her name properly. Ask for her name.’ (Possible answer: What’s her name?)

Additionally, we included 12 filler trials with neutral contexts to elicit simple declarative constructions. Furthermore, six practice items were constructed. They consisted of similar contexts, evoking CQs, NEQs and declaratives. The order of the contexts was pseudorandomized

and no more than two contexts of the same condition occurred in a sequence. Furthermore, we switched the overall order of the contexts after six speakers to avoid possible position effects.

2.1.3. Procedure

In the *Discourse Completion Task*, participants were asked to spontaneously answer to a given situation, e.g., the contexts illustrated in (2) and (3). The contexts were presented by the first author for the recording sessions in Konstanz and by the fourth author for the sessions in Frankfurt (see below). Each session started with a short practice phase in order to familiarize the participants with the procedure. Participants were asked to produce their (semi-)spontaneous reactions as natural as possible. Feedback was provided only in cases of hesitations or complete misunderstandings. In such cases, participants were asked to repeat the utterance.

The sessions of seven of the participants took place in a sound-attenuated cabin at the *PhonLab* of the University of Konstanz. Productions were recorded using an MXL 990 condenser microphone and a Tascam HD P2 portable stereo audio recorder. In addition, two participants were recorded in a sound-attenuated booth at the phonetics laboratory at the Goethe University Frankfurt using an Edirol by Roland, R-44 solid-state four-channel portable field recorder. A sampling rate of 44.1 kHz and a resolution of 16 Bit was used in all recordings.

2.1.4. Data treatment & analysis

Overall, we recorded 248 CQs and NEQs. 59 items had to be excluded due to glottalization or infelicitous answers (e.g., the context was not understood properly even after repetition). The remaining 189 productions (CQ: 91; NEQ: 98) were analyzed phonologically and phonetically using *Praat* [15]. The files were labeled segmentally at the level of morphological word and syllable. Pitch accents and boundary tones were labeled following the guidelines of Sp_ToBI presented in [16]. Pitch values and the calculated pitch range in semitones (st) were extracted automatically using a *Praat* script. Pitch was extracted within the range of the nuclear configuration at manually set minimum and maximum points.

2.2. Results

2.2.1. Intonation contours

The results across all speakers and regions show that both CQs and NEQs are predominantly realized with a rising nuclear configuration consisting of a low nuclear accent followed by a high boundary tone (L* H%; CQ: 57%, N=52; NEQ: 46%, N=45). For examples of rising contours, see Figure 1, left and mid panel. The second most common contour across both question types displays a low nuclear accent followed by a downstepped high boundary tone (L* !H%; CQ: 6%, N=5; NEQ: 17%, N=17). The third and fourth most common nuclear configurations are L* L% (CQ: 9%, N=8; NEQ: 7%, N=7) and H* L% (CQ: 8%, N=7; NEQ: 6%, N=6) respectively. Moreover, we find additional tonal configurations. The percentage distribution of the attested nuclear contours is shown in Figure 2. With respect to the attested boundary tones, our data show a predominant use of H% (CQ: 65%, N=59; NEQ: 50%, N=49). The second most common boundary tone is !H% (CQ: 14%, N=13; NEQ: 27%,

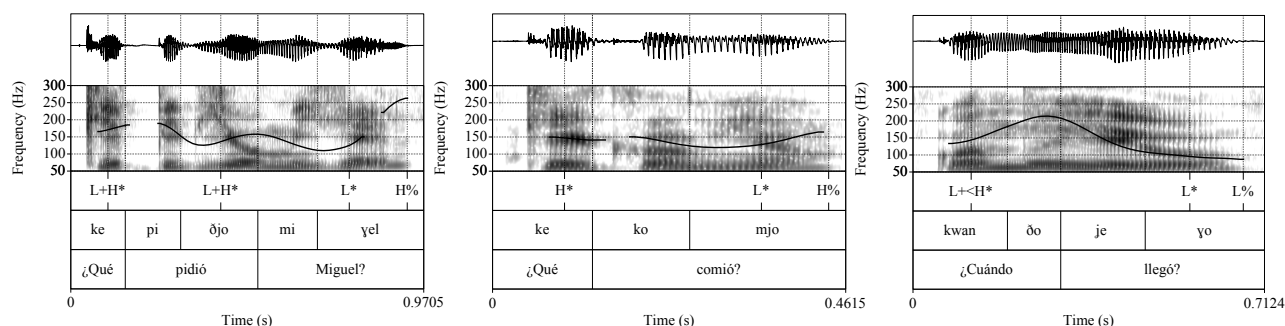


Figure 1: Waveforms, spectrograms and smoothed F0-contours. Left: CQ *¿Qué pidió Miguel?* 'What ordered Miguel?' (rising nuclear contour, male speaker, highlands). Mid: NEQ *¿Qué comió?* 'What did s/he eat?' (rising nuclear contour, same speaker as before). Right: CQ *¿Cuándo llegó?* 'When did s/he arrive?' (low nuclear contour, male speaker, lowlands).

N=26) followed in number of productions by L% (CQ: 19%, N=17; NEQ: 17%, N=17). The falling boundary tone HL% occurs in CQs only in 2% (N=2) and in NEQs in 6% (N=6).

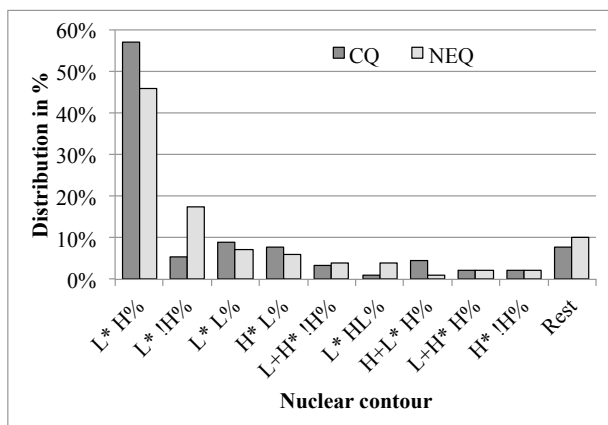


Figure 2: Distribution of the nuclear configurations in CQs and NEQs across all speakers and regions.

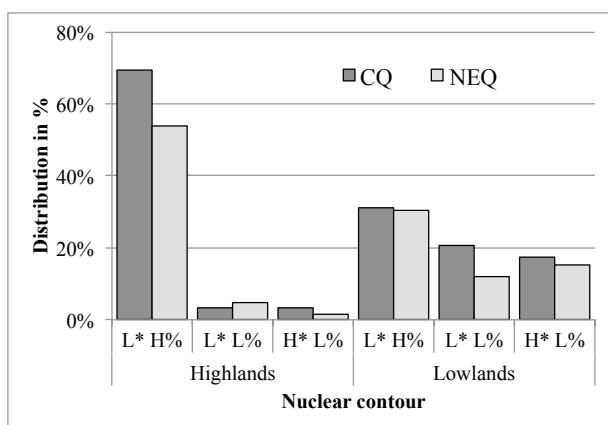


Figure 3: Distribution of rising (L* H%), low (L* L%) and falling (H* L%) nuclear contours in CQs and NEQs in highlands and lowlands.

Beyond that, the data reveal diatopic differences in the realization of the nuclear configurations. Overall we analyzed 127 utterances from highland speakers (CQ: 62; NEQ: 65) and 62 utterances from lowland speakers (CQ: 29; NEQ: 33).

While L* H% is by far the most common configuration in the productions by the highland speakers (CQ: 69%, N=43; NEQ: 54%, N=35), it is less so for the lowland speakers (CQ: 31%, N=9; NEQ: 30%, N=10). Next to L* H%, speakers from the lowlands also realize L* L% and H* L% with a comparatively high frequency (L* L%; CQ: 21%, N=6; NEQ: 12%, N=4 [see Figure 1, right panel] and H* L%; CQ: 17%, N=5; NEQ: 15%, N=5). By comparison, speakers from the highlands produce L* L% in 3% (N=2) of the CQs and in 5% (N=3) of the NEQs and H* L% in 3% (N=2) of the CQs and in 2% (N=1) of the NEQs (see Figure 3).

2.2.2. Pitch range and mean pitch

No clear tonal differences between the intonation contours of the two question types investigated are found. However, an additional analysis of the pitch range (measured in semitones, st) within the nuclear configuration and the mean pitch throughout the questions reveals acoustic differences between CQs and NEQs.

Pitch range measured in the nuclear configuration was statistically analyzed using a linear mixed effects regression model (R-version 3.1.2) with *question type* (CQ vs. NEQ) as fixed factor and *participants* and *items* as crossed random factors, allowing for random adjustments of intercepts and slopes [17]. To ensure the validity of the model, data points with residuals beyond 2.5 standard deviations were removed and the model was refitted. P-values were calculated using the Satterthwaite approximation in the R-package lmerTest.

Results for pitch range of the nuclear configuration show a significant main effect of *question type*. Participants realize CQs with larger pitch ranges than NEQs ($\beta=2.03$, $SE=0.39$, $t=5.19$, $p<0.00001$). The difference in the nuclear configuration of CQs across all speakers is on average 2.03 st larger than in NEQs (CQ: 8.96 st vs. NEQ: 6.93 st), as visualized in Figure 4 for L* H% contours. Regarding the geographical difference (highlands vs. lowlands), both speaker groups show this difference in pitch range. Speakers from the highlands produce the CQs with a 1.98 st wider pitch range than the NEQs ($\beta=1.98$, $SE=0.42$, $t=4.68$, $p<0.00001$, CQ: 9.15 st vs. NEQ: 7.17 st). Speakers from the coastal region show a 2.09 st larger pitch range in NEQs ($\beta=2.09$, $SE=0.80$, $t=2.61$, $p<0.05$) than in CQs (CQ: 8.55 st vs. NEQ: 6.46 st).

Differences in the mean pitch throughout the utterances were also statistically analyzed using a linear mixed effects regression model (R-version 3.1.2) with the same specification

and procedure as above (with *question type* as fixed factor and *participants* and *items* as crossed random factors). Again, data points with residuals beyond 2.5 standard deviations were removed and p-values were calculated using the Satterthwaite approximation in the R-package lmerTest.

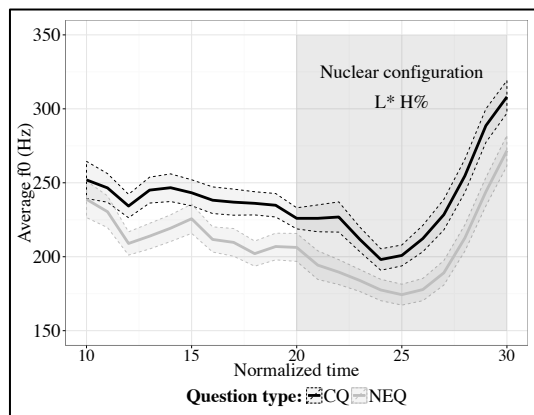


Figure 4: Time normalized F0-contour averaged across male and female speakers and both regions of the most frequent nuclear configuration (L* H%) in CQs and NEQs.

Results for mean pitch show a significant main effect of *question type*, as well. Productions of CQs show a higher F0-mean than those of NEQs ($\beta=14.95$, $SE=3.13$, $t=4.78$, $p<0.00001$). The mean pitch in CQs is on average 15 Hz higher than in NEQs (CQ: 210 Hz vs. NEQ: 195 Hz). However, in a further analysis, we ran separate models for participants from the highlands and the lowlands, to account for possible diatopic effects and found that the main effect of *question type* on the produced mean F0 within the utterances only holds true for participants who originated from the highlands ($\beta=21.13$, $SE=3.27$, $t=6.46$, $p<0.00001$). This speaker group produces CQs with an average mean pitch of 264 Hz, whereas the NEQs show a mean pitch of 243 Hz. There is no such effect of *question type* in the productions of participants from the lowlands ($\beta=2.81$, $SE=6.34$, $t=0.44$, $p=0.6$). The difference in the mean pitch of the productions of lowland speakers is only 3 Hz (CQ: 149 Hz vs. NEQ: 146 Hz).

3. Discussion

The results of the production experiment reveal that H1 can only partly be verified. In accordance with [11] we find the same nuclear contour in the productions of both NEQs and CQs. However, our results do not show a rising-falling boundary tone (HL%), as claimed by O'Rourke [11], but a rising boundary tone with a preceding low pitch accent (L* H%). Additionally, 17% of the neutral questions end in a shallower version of this contour, namely in a high plateau (L* !H%), which may be due to the narrower pitch range expansion across the nuclear contour. This suggests that incredulity is not expressed by means of tonal differences in Ecuadorian Spanish.

Interestingly, we find a difference in the tonal realization of the nuclear configuration between speakers from the highlands and the lowlands. The group from the highlands prefers L* H%, while the lowland group also frequently produces L* L% and H* L% next to L* H%. This result can be interpreted as possible intonational support for the well-

known division between *tierras bajas* ('lowlands') and *tierras altas* ('highlands') - terms that are used to capture the dialectal phonological variation on the segmental level of Latin American Spanish [18].

Regarding H2, our results support the findings of [11] that a difference between CQs and NEQs is established by a wider pitch expansion within the nuclear contour of CQs. Our analysis reveals that the pitch range difference between CQs and NEQs is significant for speakers of both geographical regions. This is in line with findings from other Romance languages that equally mark non-neutral questions by a wider pitch range ([9], [10], [19] and [20]). As a basic principle, the intonational meaning of CQs is conveyed acoustically, through differences in the pitch range of the nuclear configuration.

As for the mean pitch, our data show that only speakers from the Andean highlands produce CQs with a significantly higher mean pitch than NEQs. Speakers from the coastal region do not use mean pitch to differentiate between CQs and NEQs. This can be interpreted as a further diatopic difference between the highlands and the lowlands of Ecuador (along with the differences in the tonal realization of the nuclear configuration). However, these findings also might be due to a gender effect: five out of the six participants from the highlands were female and all speakers from the lowlands were male. Therefore, we cannot exclude a possible effect of gender rather than region. Nevertheless, differences in mean pitch seem to be a possible factor for speakers to go back to in order to discriminate between CQs and NEQs, signaling incredulity by raising the pitch level.

4. Summary and Outlook

In this study, we investigated whether neutral echo *wh*-questions (NEQs) and counterexpectational echo *wh*-questions (CQs) in Ecuadorian Spanish can be distinguished by intonational means. Our results suggest that speakers do not differentiate between CQs and NEQs by means of tonal movement. Both question types were mainly realized with a low-rising contour (L* H%), with some diatopic differences that we interpret as an intonational reflection of the classical diatopic division of Latin American Spanish in highlands and lowlands. The finding of a low-rising contour, which is also attested for CQs in other Romance varieties (e.g. in [6] and [8]), contradicts observations made by [11], who found that a rising-falling contour is used to mark those types of questions.

While no tonal differences can be attested in the data, our results show a clear distinction regarding the pitch range in the nuclear contour. This suggests that speakers of Ecuadorian Spanish (highlands and lowlands) express the difference between CQs and NEQs by means of pitch range expansion. Additionally, speakers may use a higher mean pitch throughout the utterance in order to express incredulity. However, a significant difference in the mean pitch is confirmed only for the Andean region by our data. For future work, it will be conducive to scrutinize a data set that carefully controls for sex and origin of the participants. This allows to carefully examine whether mean pitch differences in the realization of CQs and NEQs result from gender effects or differences in the geographical region. In addition, a higher number of participants may be considered. Furthermore, a perception experiment will help to show whether the differences in pitch range and mean pitch are a salient factor for hearers to discriminate between CQs and NEQs.

5. References

- [1] E.-J. Noh, "Echo questions: Metarepresentation and pragmatic enrichment," *Linguistics and Philosophy*, vol. 21, pp. 603-628, 1998.
- [2] V. Escandell-Vidal, "Intonation and procedural encoding: The case of Spanish interrogatives," in *Current Issues in Relevance Theory*, V. Rouchota and A. H. Jucker, Eds., Amsterdam: John Benjamins, 1998, pp. 169-203.
- [3] V. Escandell-Vidal, "Echo-syntax and metarepresentations," *Lingua*, vol. 112, pp. 871-900, 2002.
- [4] V. Escandell-Vidal, "Los enunciados interrogativos: Aspectos semánticos y pragmáticos," in *Gramática Descriptiva de la Lengua Española*, vol. 3, I. Bosque and V. Demonte, Eds., Madrid: Espasa Calpe, 1999, pp. 3929-3991.
- [5] C. de-la-Mota, P. M. Butragueño, and P. Prieto, "Mexican Spanish intonation," in *Transcription of Intonation of the Spanish Language*, P. Prieto and P. Roseano, Eds., München: Lincom Europa, 2010, pp. 319-350.
- [6] L. Astruc, E. Mora, and S. Rew, "Venezuelan Andean Spanish intonation," in *Transcription of Intonation of the Spanish Language*, P. Prieto and P. Roseano, Eds., München: Lincom Europa, 2010, pp. 191-226.
- [7] C. Gabriel, I. Feldhausen, A. Pešková, L. Colantoni, S.-A. Lee, V. Arana, and L. Labastía, "Argentinian Spanish intonation," in *Transcription of Intonation of the Spanish Language*, P. Prieto and P. Roseano, Eds., München: Lincom Europa, 2010, pp. 285-318.
- [8] E. Estebas-Vilaplana and P. Prieto, "Castilian Spanish intonation," in *Transcription of Intonation of the Spanish Language*, P. Prieto and P. Roseano, Eds., München: Lincom Europa, 2010, pp. 17-48.
- [9] B. Gili Fivela, C. Avesani, M. Barone, G. Bocci, C. Crocco, M. D'Imperio, R. Giordano, G. Marotta, M. Savino, and P. Soriano, "Intonational phonology of the regional varieties of Italian," in *Intonation in Romance*, S. Frota and P. Prieto, Eds., Oxford: Oxford University Press, 2015, pp. 140-197.
- [10] P. Prieto, J. Borràs-Comes, T. Cabré, V. Crespo-Sendra, I. Mascaró, P. Roseano, R. Sichel-Bazin, and M. Vanrell Bosch, "Intonational phonology of Catalan and its dialectal varieties," in *Intonation in Romance*, S. Frota and P. Prieto, Eds., Oxford: Oxford University Press, 2015, pp. 9-62.
- [11] E. O'Rourke, "Ecuadorian Andean Spanish intonation," in *Transcription of Intonation of the Spanish Language*, P. Prieto and P. Roseano, Eds., München: Lincom Europa, 2010, pp. 227-253.
- [12] S. Blum-Kulka, J. House, and G. Kasper, "Investigating cross-cultural pragmatics: An introductory overview," in *Cross-Cultural Pragmatics: Requests and Apologies*, S. Blum-Kulka, J. House, and G. Kasper, Eds., Norwood: Ablex, 1989, pp. 1-34.
- [13] P. Prieto and P. Roseano, Eds., *Transcription of Intonation of the Spanish Language*. München: Lincom Europa, 2010.
- [14] P. Prieto and P. Roseano. (2009-2010). *Atlas interactivo de la entonación del español*. Available: <http://prosodia.upf.edu/atlasentonacion/>
- [15] P. Boersma and D. Weenink, "Praat: doing phonetics by computer," 5.3.14, 1992-2015.
- [16] J. I. Hualde and P. Prieto, "Intonational variation in Spanish: European and American varieties," in *Intonational Variation in Romance*, S. Frota and P. Prieto, Eds., Oxford: Oxford University Press, 2015, pp. 350-391.
- [17] D. J. Barr, R. Levy, C. Scheepers, and H. J. Tily, "Random effects structure for confirmatory hypothesis testing: Keep it maximal," *Journal of Memory and Language*, vol. 68, pp. 255-278, 2013.
- [18] J. M. Lipski, *Latin American Spanish*. London: Longman, 1994.
- [19] S. A. Lee, F. Martínez-Gil, and M. E. Beckman, "The intonational expression of incredulity in absolute interrogatives in Buenos Aires Spanish," in *Selected Proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology*, M. Ortega-Llebaria, Ed., Somerville, MA: Cascadia, 2008, pp. 47-56.
- [20] V. Crespo-Sendra, M. d. M. Vanrell, and P. Prieto, "Information-seeking questions and incredulity questions: Gradient or categorical contrast," in *Proceedings of the 5th International Conference on Speech Prosody, May 11-14, Chicago, USA, Proceedings*, 2010, pp. 11-14.