ISCA Archive
http://www.isca-speech.org/archive

Speech Prosody 2010
Chicago, IL, USA
May 10-14, 2010

Nuclear rises and final rises in Manchego Peninsular Spanish yes/no questions

Nicholas C. Henriksen

Department of Linguistics & Department of Spanish & Portuguese
Indiana University, USA
nhenriks@indiana.edu

Abstract

This paper is an experimental investigation examining the tonal structure of yes/no question intonation by speakers of Manchego Peninsular Spanish. It provides a phonetic and phonological analysis of a corpus of 738 yes/no question utterances produced by 16 speakers in a contextualized sentence reading task. The acoustic-phonetic analysis focuses on the scaling and timing correlates of final rises produced under various tonal clash and non-clash contexts. The quantitative results provide evidence for two separate tonal configurations, and this difference is indicated by contrasting nuclear pitch accent specifications: H*…¡H% and L*…H%.

Index Terms: yes/no questions, final rises, sparse tonal specification

1. Introduction

In keeping with the Autosegmental-metrical framework to intonational phonology, it is assumed that utterance melodies are the realization of abstract sequences of H(igh) and L(ow) tones [2,7]. These tones are organized into pitch accents, which associate to metrically strong (i.e., stressed) syllables, and boundary tones, which associate to utterance edges. Phonetically, tonal targets are defined by local F0 minima or maxima turning points quantified on two dimensions: scaling, or F0 value; and timing, or relative alignment to a segmental boundary. Crucially, AM adopts the notion of *sparse tonal specification* [1] in assuming that the factors that determine F0 may involve nothing more than interpolation from an earlier pitch target to a later one. Given that AM treats utterance melodies as strings of tonal events, not all syllables are required to bear tonal specifications. The advantage of adhering to the notion of sparse tonal specification is that it makes possible the description of contours on utterances such as *Where?* and *Where are you going?*, for example, as phonologically identical while still allowing to model the phonetic details that result from the difference in utterance length. That is, while the realization of tonal targets may be more compressed in the one-syllable utterance than in the longer utterance, the two surface productions are assumed to represent identical melodic specifications.

The current analysis is concerned with the phonetics and phonology of utterance-final rises in yes/no questions. Research on question signaling indicates that questions are typically marked by a higher pitch level than lexically identical declaratives, although the exact surface manifestation of the pitch difference varies cross-linguistically [9]. It is also well known that the final boundary rise is the most recurrent prosodic characteristic used to communicate a yes/no question. Importantly, the alignment of the boundary rise has been shown to exhibit a great deal of cross-linguistic variation, motivated by discourse-level, structural, and even sociolinguistic factors. For example, Warren & Daly [11] reported that final rises in intonation-only questions in New Zealand English begin as early as the onset of the nuclear stressed syllable or as late as the onset of posttonic material. More recently, Warren [10] uncovered that speaker age and speaker sex in addition to structural factors contribute to variable final rise alignment in questions. He noted that characterization of the utterance-final rise as a steadily rising movement from the nuclear syllable may be too restrictive.

Experimental research on Spanish yes/no question intonation indicates that its utterance melody consists of a prenuclear rise followed by nuclear valley and a boundary rise. Face [4] carried out a production experiment examining the differences between lexically identical declarative statements and yes/no questions. Yes/no questions exhibited a significantly higher prenuclear peak than declaratives, although no alignment differences could be found in the L+H rise. Utterance-medially, declaratives exhibited a rising pitch accent, whereas questions did not. Utterance-finally, the boundary movements were quite different, with a sharp rise in the yes/no questions and a fall or level tone in the declaratives. The final rise typically occurred at the onset of posttonic material. Work by Willis [12] on Dominican Spanish, however, has shown that the start of the boundary rise may occur within the bounds of the nuclear stressed syllable.

The Spanish yes/no question final boundary rise is generally assumed to reflect a H% specification, due in great part to its late (i.e., posttonic) alignment. Although Willis [12] reported some variability in the Dominican Spanish final rise, the phonetic and/or phonological implications of such variable alignment have not been addressed. What remains to be seen is whether final rise alignments in Spanish may be different because they are phonetic implementations of different tonal targets, or because phonetic implementation rules operate on the same phonological representation in different conditions. The current investigation seeks to answer exactly this question by examining a corpus of 738 yes/no questions tokens which do not seem to demonstrate consistent final rise alignment on initial inspection. In doing so, it recognizes the notion of sparse tonal specification as fundamental for understanding the linguistic expression of underlying tonal melodies. This is reflected in the research design and the construction of the target utterances used in the sentence reading task.

2. Procedure

2.1. Participants and data elicitation

A total of 16 native speakers (8 male; 8 female) of Manchego Peninsular Spanish participated in a contextualized sentence reading task designed to elicit broad focus declarative, yes/no question, and wh-question utterances. Speaker responses to yes/no question prompts only are analyzed in the current study. The sentence reading task consisted of a 2-click PowerPoint contextualized prompt methodology. All target sentences were preceded by a discourse prompt so that the pragmatic intent of the target utterance would be clear. Each target utterance contained two lexical words, and combinations of lexical words were created based on two manipulation conditions: number of unstressed syllables between the prenuclear and nuclear stressed syllables (0-3); and number of unstressed syllables after the nuclear stressed syllable (0-2). This yielded a total of 12 manipulation contexts. Two lexical combinations were created per manipulation context, and each lexical combination was uttered twice, for a total of 48 target yes/no question productions per speaker (12x2x2=48). The data were recorded using a SONY HI-MD MZ-RH1 minidisc recorder and a Shure WH20 head-mounted microphone. The task lasted approximately 16-20 minutes, and participants were recorded in a quiet room by the researcher in Socuéllamos (Ciudad Real), Spain.

2.2. Data analysis

Duration and F0 measurements were obtained for the target productions using Praat [3]. Speaker productions were manually segmented and labeled by simultaneous inspection of waveforms and wide-band spectrograms following standard segmentation criteria. F0 points thought to represent tonal targets were also located manually. A Low point was defined as an 'elbow' where a low level stretch turned into a clear rise. A High point was the highest F0 value of a peak from rise-fall movement. Frequency values were calculated in the ERB scale. A total of 30 productions were discarded from the original corpus of 768 yes/no questions. Tokens were discarded if they contained hesitation pauses, disfluencies, or for reasons to be detailed in the following section.

3. Results

Initial inspection of the yes/no question corpus revealed two distinct final rise type patterns. The first pattern, referred to here as an 'early rise', typically began near the onset of the nuclear stressed syllable and continued its ascent throughout the remainder of the utterance. The tonal level of the F0 minimum was slightly lower than that of the preceding peak. A sample contour is provided in Figure 1 (stressed syllables are denoted by capital letters). The second pattern, referred to here as a 'late rise', typically began near the end of the nuclear stressed syllable and in some cases began within the bounds of posttonic segmental material. The tonal level of the F0 minimum was considerably lower than that of the prenuclear peak, and in most cases, it approximated the F0 of the initial tonal value. The tonal descent from the prenuclear peak to the start of the final rise was greater in late rise contours than in early rise contours. A sample late rise contour is provided in Figure 2.

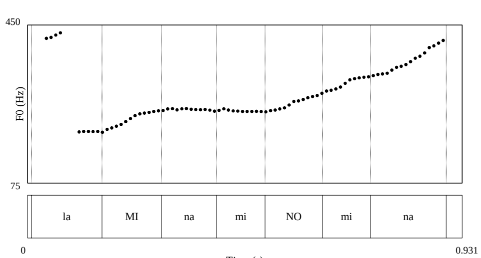

Figure 1: *Sample early rise production.*

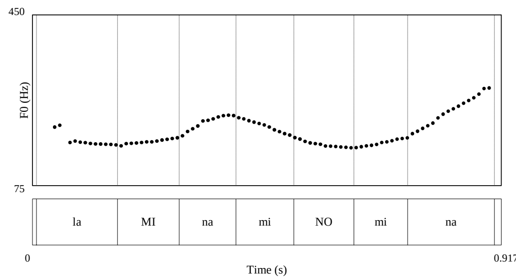

Figure 2: *Sample late rise production.*

3.1. Relative rise start in non-clash and final clash contexts

A number of acoustic measurements were taken to provide an objective measure of the timing and scaling differences in the two rise type patterns. Tokens were also grouped according to prosodic manipulation context: internal clash (i.e., where no unstressed syllables separated the prenuclear and nuclear stressed syllables); final clash (i.e., where the nuclear stressed syllable was the utterance-final syllable); and non-clash.

A relative rise start score was obtained for each final rise token by diving the temporal distance of the rise start (i.e., the F0 minimum) from the beginning of the nuclear stressed syllable by the total syllable duration. A rise start aligned exactly at the stressed syllable onset would receive a score of '0', and a rise start aligned exactly at the stressed syllable offset would receive a score of '100'. The data in Figure 3 present a distributional analysis of the relative rise start scores for each speaker in non-clash contexts. As can be seen, six speakers (6F, 7F, 8F, 6M, 7M, 8M) produced final rises that never began beyond the 40% point of the nuclear stressed syllable. Their rises typically began within the first third of that syllable and in some cases began before the syllable onset. The remaining ten speakers, however, produced rises that typically began within the second half of the nuclear stressed syllable. For most speakers, the median rise start was between 90% and 100%, although one speaker (5M) obtained a median score of approximately 50%.

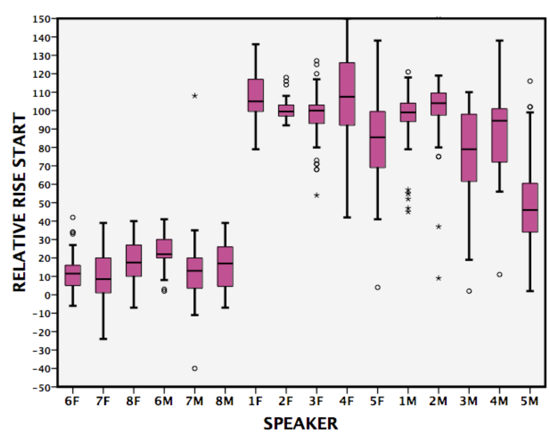

Figure 3: *Relative rise start scores (non-clash contexts).*

Given the distributional data in Figure 3, tokens produced by speakers 6F, 7F, 8F, 6M, 7M, and 8M were labeled as 'early rises' and tokens produced by the remaining ten speakers were labeled as 'late rises'. Each set of data was then submitted to statistical analysis in which early rise productions were compared to late rise productions. In order to ensure direct comparability, positive outliers for early rise speakers and negative outliers for late rise speakers were discarded.

	NON-CLASH			FINAL CLASH		
SPKR.	**Rise start**	**Syll dur**	**RRS**	**Rise start**	**Syll dur**	**RRS**
6F	33ms	135ms	15%	23ms	217ms	10%
7F	18ms	132ms	14%	0ms	176ms	0%
8F	27ms	133ms	21%	11ms	175ms	6%
6M	36ms	158ms	14%	40ms	212ms	14%
7M	14ms	112ms	11%	0ms	183ms	1%
8M	13ms	118ms	10%	29ms	227ms	13%
AVG(ER)	**21ms**	**126ms**	**16%**	**12ms**	**186ms**	**6%**
1F	149ms	141ms	106%	100ms	211ms	47%
2F	137ms	137ms	100%	89ms	205ms	43%
3F	141ms	155ms	90%	96ms	259ms	37%
4F	118ms	118ms	100%	61ms	171ms	35%
5F	119ms	155ms	77%	44ms	221ms	25%
1M	130ms	142ms	89%	51ms	131ms	23%
2M	124ms	128ms	97%	84ms	229ms	37%
3M	95ms	129ms	73%	51ms	200ms	24%
4M	100ms	124ms	81%	64ms	173ms	37%
5M	50ms	127ms	38%	29ms	227ms	17%
AVG(LR)	**115ms**	**135ms**	**85%**	**67ms**	**204ms**	**32%**
ER-LR	**−94ms**	**−9ms**	**−69%**	**−55ms**	**−18ms**	**−26%**
SIG.	***p<.001***	***p<.001***	***p<.001***	***p<.001***	***p<.001***	***p<.001***

Table 1: *Final rise alignment.*

The data in Table 1 provide information for rise start scores in non-clash and final clash contexts. As can be seen, relative rise start scores for early rise speakers in non-clash contexts averaged to 16%. For late rise speakers in non-clash contexts, the average score was 85%. This difference was shown to be significant. A second analysis examined whether the early vs. late rise contrast held in cases of final clash. Given that no posttonic material is available in these contexts, it may be hypothesized that all speakers realize an 'early rise' (i.e., towards the beginning of the stressed syllable) to allow for tonal ascent and realization of the final tonal level. The data in Table 1 show that this was not the case. Although rise start scores for both speaker groups were earlier in final clash contexts as compared to non-clash contexts, the difference between early rise speakers (6%) and late rise speakers (32%) still proved significant. What this indicates is that even in cases of extreme tonal pressure of an upcoming boundary tone, the rise start difference holds.

SPKR.	**IT**	**PH**	**RS**	**FT**
6F	5.79	7.00	6.81	9.45
7F	5.45	6.31	6.16	8.26
8F	5.66	6.78	6.78	8.65
6M	4.77	5.93	5.70	7.69
7M	3.55	4.62	4.58	5.95
8M	3.82	5.58	5.50	8.19
AVG(ER)	**5.21**	**6.04**	**5.92**	**8.03**
1F	5.00	6.31	4.76	7.29
2F	5.04	6.18	4.83	7.24
3F	5.70	6.61	5.56	7.99
4F	6.21	6.76	5.95	8.61
5F	6.12	7.41	6.01	9.12
1M	4.03	5.29	4.19	6.57
2M	3.49	5.21	3.49	6.59
3M	3.63	4.58	3.72	6.72
4M	3.48	4.13	3.55	4.89
5M	3.42	4.65	3.82	6.61
AVG(LR)	**4.69**	**5.71**	**4.59**	**7.16**
ER-LR	**0.23**	**0.32**	**1.33**	**0.87**
SIG.	***p>.05***	***p>.05***	***p<.001***	***p<.001***

Table 2: *F0 measurements (non-clash and final clash).*

Next, F0 measurements were taken to determine possible scaling differences between the two rise types. The data in Table 2 provide information for the F0 of the initial tonal value (IT), the prenuclear peak (PH), the rise start (RS), and the final tonal value (FT). As can be seen, scaling differences between the two groups did not prove significant for the utterance-initial value and the prenuclear peak. A significant effect was found for speaker group on rise start value and on final tonal value, however, and in both cases, early rises exhibited higher F0. Importantly, early rises began at approximately the same tonal level of the prenuclear peak (5.92 vs. 6.04 ERB, respectively), whereas late rises began at approximately the same tonal level of the initial tonal value (4.59 and 4.69 ERB, respectively).

3.2. Internal clash contexts

This section reports on data extracted from question tokens produced in internal clash contexts (i.e., where no unstressed syllables separated the prenuclear and nuclear stressed syllables). Initial inspection of speaker data in these contexts revealed that the contrast between speaker groups was not limited to the relative timing or scaling of the final rise, but to the overall shape or configuration of the contour itself. A sample production from an early rise speaker is provided in Figure 4, and a sample production from a late rise speaker is provided in Figure 5. As can be seen, the contour produced by the early rise speaker contains a continuous F0 uptrend with no indication of tonal descent, whereas the contour produced by the late rise speaker consists of a rise-fall movement across the two stressed syllables. Crucially, the six speakers who produced 'early rises' in non-clash contexts categorically produced contours similar to the one in Figure 4 in internal clash contexts. The ten speakers who produced 'late rises' in non-clash contexts categorically produced contours similar to the one in Figure 5 in internal clash contexts.

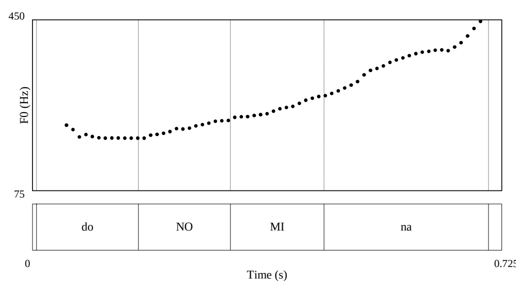

Figure 4: *Sample internal clash production (early rise).*

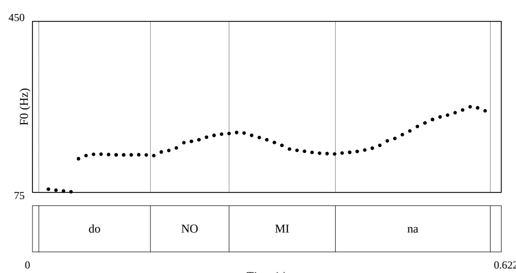

Figure 5: *Sample internal clash production (late rise).*

In order to provide an objective measure of the differences in contour production in internal clash contexts between the two speaker groups, F0 measurements were taken at the beginning (i.e., C1) and the end (i.e., C2) of the nuclear stressed syllable. The pitch trajectory throughout this syllable was then calculated by subtracting the F0 at the syllable onset from the F0 at the syllable offset, and these data are provided in Table 3. As can be seen, all early rise speaker productions exhibited a rising pitch trajectory throughout the nuclear stressed syllable, and all late rise speakers exhibited a falling trajectory throughout the nuclear stressed syllable. Early rise speakers rose 0.64 ERB on average, whereas late rise speakers fell 0.92 ERB on average. This difference was significant.

SPKR.	ERB (C1)	ERB (C2)	DIFF (C1-C2)
6F	6.47	7.46	1.00
7F	3.86	4.49	0.62
8F	6.38	6.90	0.52
6M	5.86	6.47	0.61
7M	6.26	6.90	0.64
8M	4.10	4.56	0.46
AVG(ER)	**5.49**	**6.13**	**0.64**
1F	6.20	4.61	-1.59
2F	5.95	4.90	-1.04
3F	6.36	5.73	-0.64
4F	6.51	5.84	-0.67
5F	7.45	6.47	-0.98
1M	5.46	3.84	-1.62
2M	4.39	3.43	-0.96
3M	4.88	3.86	-1.01
4M	3.92	3.62	-0.30
5M	4.61	4.18	-0.43
AVG(LR)	**5.57**	**4.65**	**-0.92**
ER-LR	**-0.08**	**1.48**	**1.56**
SIG.	***p>.05***	***p<.001***	***p<.001***

Table 3: *F0 in internal clash contexts.*

4. Discussion and conclusion

The acoustic-phonetic analysis has brought to light a number of differences between the two speaker groups of Manchego Peninsular Spanish under examination here. First, significant timing differences were exhibited between the speaker groups in non-clash and final clash contexts. In non-clash contexts, the early rise speakers never surpassed the 40% threshold, whereas the late rise speakers realized the F0 minimum within the second half of the nuclear stressed syllable, and in many cases the rise start was coincident with the onset of posttonic material. The rise start contrast was shown to hold in final clash contexts as well, although the timing differences were compressed. Regarding scaling differences, early rises began at approximately the same level of the previous tonal peak, whereas late rises began at approximately the same level of the initial tonal value. Early rises were also shown to have a higher final tonal value. Finally, production data in internal clash contexts were helpful for discerning the tonal specification of each rise type. Early rise speakers rose throughout the nuclear stressed syllable when no unstressed syllables separated it from the prenuclear syllable, indicative of a nuclear High tone specification in addition to the final boundary rise. Late rise speakers, however, produced a rise-fall-rise, indicative of a Low target positioned between two Highs.

In terms of phonological implications, the early rise and late rise have been shown to differ on the two phonetic dimensions known to characterize tonal targets within the AM framework: F0 level and temporal coordination with the segmental string. The timing contrast was shown to hold even in cases of final tonal clash where the lack of posttonic segmental material might be expected to obscure the early vs. late rise start time difference. Another indication that different tonal specifications triggered the different rise starts was found in speaker responses to internal clash. Response differences in tonal clash contexts are typically assumed to indicate differences in tonal specifications [8]; thus it can be inferred that different phonological targets triggered the upward vs. rise-fall-rise movements in internal clash contexts in the current corpus. We express the difference by means of contrasting nuclear pitch accent specifications: H*…H% for early rises and L*…H% for late rises. Additionally, higher FT in the early rise contour is denoted by upstepped ¡H%, thus H*…¡H% (vs. L*…H% for the late rise). In keeping with the notion of sparse tonal specification, we assume that the tonal plateau or the slight descent from the prenuclear peak to the F0 minimum in the early rise was a consequence of phonetic interpolation, or F0 sag between two H tone specifications [6,7]. Crucially, no evidence could be found for a sagging transition in internal clash utterances. The tonal descent exhibited in late rise productions, however, and most notably in internal clash contexts, provides evidence for a nuclear L* specification between the prenuclear H and the final H%.

In sum, research on Spanish yes/no question intonation can no longer assume that utterance-final rises may only be triggered by a H% boundary tone specification. As the current investigation has shown, two final rise options are available. On the one hand, a nuclear H* may pull the rise start toward the beginning of the stressed syllable. This is followed by a ¡H% boundary which accounts for the sustained pitch rise throughout the remainder of the utterance. On the other hand, a nuclear L* may trigger a sharp tonal descent from the previous peak. This is followed by a later final rise start time. Although the rise starting point in the latter may begin within the bounds of the nuclear stressed syllable, it is considerably later than the early rise and is not compressed in cases of internal tonal clash, providing evidence for H% in addition to nuclear L*. On a final note, the current paper has not addressed the pragmatic and/or stylistic meanings of the two rise types under investigation. For elaboration on this topic, the reader is directed to Henriksen [5].

5. References

[1] Arvaniti, A. and Ladd, D.R. 2009. Greek wh-questions and the phonology of intonation. *Phonology 26,* 43-74.

[2] Beckman, M.E. and Pierrehumbert, J.B. 1986. Intonational structure of Japanese and English. *Phonology 3*, 255-309.

[3] Boersma, P. and Weenink, D. 2009. *PRAAT: Doing phonetics by computer, version 5.0.32.* <http://www.praat.org>.

[4] Face, T.L. 2004. The intonation of absolute interrogatives in Castilian Spanish. *Southwest Journal of Linguistics 23*, 65-79.

[5] Henriksen, N.C. In prep. *Question intonation in Manchego Peninsular Spanish*. PhD dissertation, Indiana University.

[6] Ladd, D.R. and Schepman, A. 2003. "Sagging transitions" between high pitch accents in English: experimental evidence. *Journal of Phonetics 31*(1), 81-122.

[7] Pierrehumbert, J.B. 1980. *The phonetics and phonology of English intonation*. PhD dissertation, MIT. Distributed by Indiana University Linguistics Club.

[8] Prieto, P. 2005. Stability effects in tonal clash contexts in Catalan. *Journal of Phonetics 33*(2), 215-242.

[9] van Heuven, V. and Haan, J. 2000. Phonetic correlates of statement versus question intonation in Dutch. In *Intonation: Analysis, Modelling and Technology,* A. Botinis (ed.). Dordrecht: Kluwer, pp. 119-144.

[10] Warren, P. 2005. Issues of late rising in New Zealand English: Intonational variation or intonational change?. *Language Variation and Change 17*, 209-230.

[11] Warren, P. and Daly, N. 2000. Sex as a factor in rises in New Zealand English. In *Gendered speech in social context: Perspectives from gown and town,* J. Holmes (ed.). Wellington: Victoria University Press, pp. 99-115.

[12] Willis, E.W. 2004. Dominican Spanish absolute interrogatives in broad focus. In *Laboratory approaches to Spanish phonology,* T.L. Face (ed.). Berlin: Mouton de Gruyter, pp. 61-79.