The role of pitch range in establishing intonational contrasts in Catalan

Joan Borràs-Comes,[1] *Maria del Mar Vanrell,*[1] *Pilar Prieto* [2,1]

[1] Universitat Pompeu Fabra, Barcelona, Spain
[2] ICREA, Barcelona, Spain

{joan.borras; mariadelmar.vanrell; pilar.prieto}@upf.edu

Abstract

In Catalan, the same rising nuclear pitch accent L+H* is used in three different sentence-types, namely statements, contrastive foci, and echo questions. Since the peak height of the rising pitch accent seems to indicate sentence type, we hypothesized that these three pragmatic meanings would be differentiated by pitch accent range. We undertook two identification tasks and analyzed the patterns of responses found as well as reaction times (RTs). The results of the identification tasks show that there is a contrast between the statement interpretation on the one hand (L+H*) and the contrastive foci and echo question interpretation on the other (L+¡H*). However, RTs clearly show that while there is a categorical difference between the statement interpretation (L+H*) and the echo question interpretation (L+¡H*), the difference between a statement interpretation and a contrastive focus interpretation is gradient. This represents further evidence that pitch range can be used to make phonological distinctions between a variety of pragmatic meanings, and strengthens the argument that this needs to be represented descriptively at the phonological level.

Index Terms: echo question intonation, contrastive focus, statement intonation, categorical perception, reaction time measures, pitch range differences, Catalan language

1. Introduction

In Catalan, the same rising nuclear pitch accent L+H* is used in three different sentence-types, namely statements (see 1a), contrastive foci (1b), and echo questions (where the speaker challenges what has been said) (1c).

(1) a. — Com la vols, la cullera?
What type of spoon do you want?
— **Petita**[, sisplau].
[I want a] little [spoon, please].

b. — No volies una cullera gran?
Didn't you want a big spoon?
— **PETITA**[, la vull, i no gran].
[I want a] little [one, and not a big one].

c. — Jo la vull petita, la cullera
I want a little spoon.
— **Petita**?[, n'estàs segur?]
[A] little [one]? [Are you sure?]

In previous analyses of Catalan dialectal data coming from the *Interactive Atlas of Catalan Intonation* [1] (see also [2]) we observed that intonation contours of these three sentence-types differ in their pitch accent height. While the rising pitch accent of statements (L+H*) is produced with a narrow pitch range (see Fig. 1, upper panel), that of contrastive foci and echo questions is produced with a much wider pitch range (see Fig. 1, bottom panel).

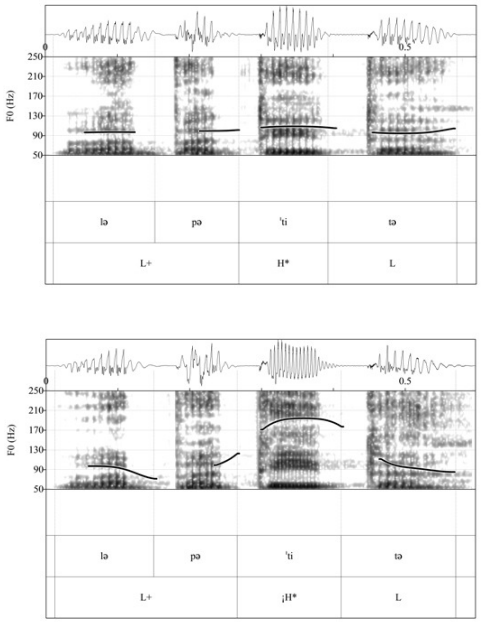

Figure 1. *Waveforms, f0 contours, and Cat_ToBI transcription of the utterance La petita 'The small one' produced with a neutral statement meaning (upper panel) and a contrastive focus or echo question meaning (bottom panel).*

However, taking statement sentences as our starting point, our perception as native speakers is that the rising pitch accent of a statement does not require much augmentation for it be understood as a contrastive focus. This leads us to hypothesize that these three sentence-types may be distributed in three well-differentiated areas of the pitch range (see Fig. 2).

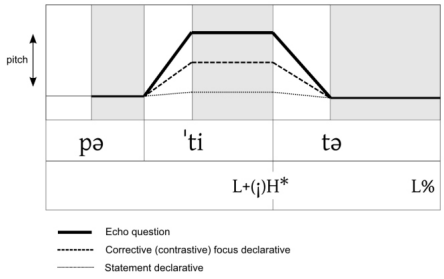

Figure 2. *Idealized intonational contours for neutral statement meaning, contrastive focus and echo question.*

Though the Autosegmental-Metrical (AM) model traditionally interpreted pitch range differences as having a gradient interpretation (namely a reflection of incremental differences in speaker involvement in the speech act) [3, 4], some recent studies have applied the paradigm of categorical perception to differences in H peak height [5, 6]. Moreover, several studies have reported categorical distinctions between pitch accents that differ in their pitch range. For example, Savino & Grice [7] reported that pitch range variation played a role in the identification of two different pragmatic meanings in Bari Italian, namely information-seeking questions and echo questions. Similarly, Vanrell [8] reported that the main cue for the categorical distinction between *wh*-questions and yes-no questions was the pitch height of the H tone of the nuclear accent in Majorcan Catalan (see also [9]).

In the abovementioned studies a two-way difference in pitch range was tested. The goal of this study is to investigate a potential three-way distinction in pitch range between a statement and contrastive focus meaning on the one hand and contrastive focus meanings and echo questions on the other. Our hypothesis is that the three types of pragmatic meanings (statement, contrastive focus, and echo question) occupy different pitch range areas, from more compressed to more expanded pitch range differences.

2. Methodology

With the goal of investigating the role of pitch range changes in the interpretation of rising pitch accents in Catalan, we undertook a series of perception experiments. Twenty native speakers participated in two semantically motivated identification tasks. The tasks dealt with the contrast between statements and echo questions (task 1) and statements and contrastive foci (task 2).

The stimuli for these tasks (see Fig. 3) consisted of a continuum that was created by modifying the F0 height of the peak in 11 steps (distance between each one = 1.2 semitones) of the noun phrase *petita* [pə.'ti.tə] (‘little’-fem). Natural productions of the two extreme contours (echo question and statement) were read by a male native speaker of Catalan, and these utterances served as the source utterances for our stimuli. The speech manipulation was performed by means of Praat [10]. The original noun phrase sentence was pronounced with a rising-falling contour L+H* L%. The rising movement was realised as a 100 ms high plateau starting 30 ms after the onset of the accented syllable /'ti/, and was preceded by a low plateau for the syllable [pə] (102.4 Hz, 100 ms). The posttonic syllable [tə] was realized with low plateau (94.5 Hz, 180 ms). The pretonic and posttonic F0 levels were maintained invariable in all manipulations. The peak height continuum ranged from 105.3Hz to 208.7 Hz.

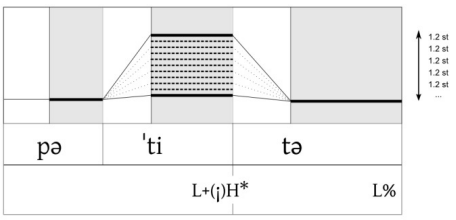

Figure 3: *Schematic contour of the pitch manipulation.*

The experiment was set up by means of the psychology software E-prime version 2.0 [11]. Seventeen native speakers of Central Catalan and three native speakers of North-western Catalan participated in the experiment. Stimuli were presented to subjects over headphones. Subjects were instructed to pay attention to the intonation of the stimuli and decide which interpretation was more likely for each stimulus by pressing the corresponding computer key. A training session was conducted prior to each task, to get subjects used to the stimuli and the tasks. The order in which the participants were asked to do the tasks was counterbalanced for half of them. Each task consisted of five blocks in which every stimulus in the continuum was presented to the subjects in a randomised order.

Chen [12] holds that the CP paradigm may be unsuitable for investigating the categoriality, but gradiency of peak height contrasts, because that paradigm relies on the inability of listeners to detect differences between two stimuli taken from the same category. She claims that the use of reaction time measures (RTs) (see, e.g. [13]) can be used in conjunction with the results of identification tasks to help interpret whether tonal perception effects are categorical or continuous. In general, mean RTs are higher when the stimulus is ambiguous and are diminished when the stimulus corresponds clearly to one category. This close correlation has been found for Majorcan Catalan [8, 9], European Portuguese [14], and Bari Italian [7].

The data of the response frequencies and RTs were automatically recorded in E-prime. A timer with 1 ms accuracy was activated at the beginning of each stimulus and the RTs were recorded from the beginning of each stimulus until a response was given. The experiment was set up in such a way that the next stimulus was presented only when a response was given; yet subjects were instructed to press the button as quickly as they could. For this reason, RTs of less than 130 ms (before the appearance of H*) or more than 5000 ms were categorized as outliers.

3. Results

3.1. Pattern of responses

Figures 4 and 5 show the identification rate for the continuum created for task 1 (Figure 4) and task 2 (Figure 5). The “identification rate” is defined as the number of “echo question” responses (Figure 4) or “focus” responses (Figure 5) over the total number of responses. The results show that both functions present a shallow S-shape. For both tasks, the response frequency-curves suggest that the first three stimuli belong to the category ‘statement’ and the last three stimuli (higher peaks) can be interpreted as either contrastive foci or echo questions, depending on the task. These results do not support the claim that echo questions and contrastive foci occupy different areas of the pitch range.

The Wilcoxon signed-rank sum test revealed that there is a significant difference between the two curves (T = 9570, p < .01, r = -.065). For both tasks, the stimuli 4, 5 and 6 appear to form the ‘dynamic zone’ of the continuum. Applying the Wilcoxon test to each pair of stimuli we found that there are two significant differences for task 1: between stimuli 4 and 5 (T = 300, p < .01, r = -.254) and between 6 and 7 (T = 67.5, p < .05, r = -.223). Yet for task 2 we only found a significant difference between stimuli 5 and 6 (T = 275, p < .01, r = -.276).

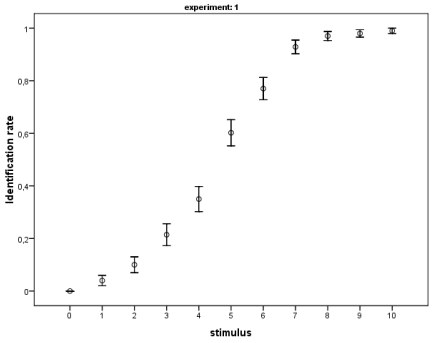

Figure 4: *Identification rate for task 1 (statement/echo).*

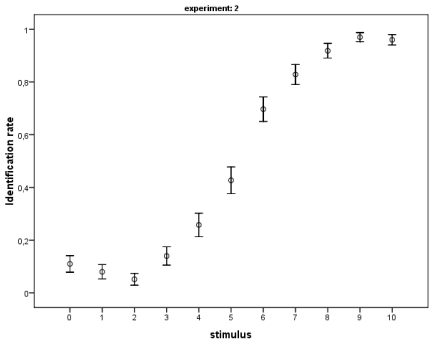

Figure 5: *Identification rate for task 2 (statement/focus).*

3.2. Reaction time measurements

As mentioned above, Reaction Time (RT) measurements have been proposed to be a good alternative to the discrimination task in testing the hypothetical discreteness of a contrast ([12], [13]). [12] (2003:100) claims that "short RTs for within-category identification and long mean RTs for across-category identification are essential properties of linguistically real identification categories". Figures 6 and 7 plot averaged RT responses (in ms) for all subjects, for both tasks (task 1: statement/echo; task 2: statement/echo question). Remember that the measures reflect RTs from the start of the utterance. The graphs indicate that even though both tasks present longer RTs for central stimuli, only in task 1 is a clear increase (and a peak) observed in the central stimuli. Wilcoxon tests revealed that there exists a significant difference between the two sets of RTs ($T = 269010$, $p < .05$, $r = -.049$).

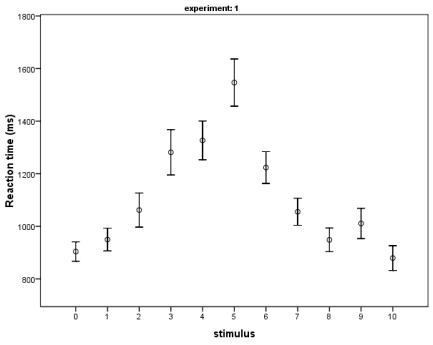

Figure 6: *Averaged reaction time measures (in ms.) for all subjects (task 1, statement/echo).*

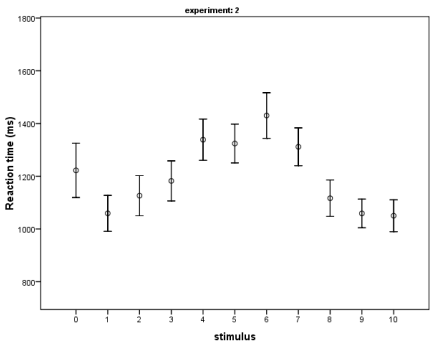

Figure 7: *Averaged reaction time measures (in ms.) for all subjects (task 2, statement/focus).*

We then compared both sets of RTs using a one-way ANOVA and Tukey HSD post-hoc tests. For both tasks, RTs differed significantly across steps (task 1: F (10, 1082) = 11.528, $p < .0001$; task 2: F (10, 1075) = 3.087, $p < .001$).

The RT graphs for task 1 show that participants were significantly slower in responding to stimuli at step 5 (and also steps 3, 4), and faster for the stimuli clearly belonging to each category. However, RTs for task 2 show a high plateau in the central stimuli instead of an RT peak, and an unexpected peak at the left extreme.

Table 1 shows the RT pairwise comparisons (asterisks indicate significant differences) for task 1. As expected, Tukey HSD post-hoc tests showed a significant difference between step 5 and all other steps in spite of 3 and 4 ($p < .05$ compared to step 6; $p < .001$ in all other comparisons). Thus we can consider step 5 to be the across-category step.

st	0	1	2	3	4	5	6	7	8	9	10
0	—			**	***	***	*				
1		—		**	**	***					
2			—			***					
3	**	**		—					*		***
4	***	**			—				**	*	***
5	***	***	***			—	*	***	***	***	***
6	*					*	—				**
7						***		—			
8				*	**	***			—		
9					*	***				—	
10				***	***	***	**				—

Table 1. *Significant differences of pairwise comparisons of task 1* (* p <.05; ** p <.01; *** $p < .001$)

By contrast, for task 2 only three significant differences obtained in all pairwise comparisons, namely when step 6 was compared with steps 1, 9 and 10 ($p < .05$ in all cases).

In summary, the response frequencies and mean RTs for task 1 show that the distinction between L+H* and L+¡H* is of a discrete nature and that it cues a statement/echo question contrast. On the other hand, the RT patterns for task 2 reveal the gradient behavior of pitch scaling in the contrast between a statement and a contrastive focus declarative. Other cues like tonal alignment have to be investigated to better disentangle this contrast.

4. Discussion and conclusions

On the one hand, our results show that there is a categorical difference between the **statement interpretation** and the **echo question interpretation**. In this case, the pitch range variation is the main cue that Catalan listeners use to decide between one and the other. This fact represents further evidence that pitch range can be used to make phonological distinctions between a variety of pragmatic meanings, and that this needs to be represented descriptively at the phonological level. In our case, Catalan intonational phonology should make a distinction between a simple rising pitch accent L+H* (for statements) and an upstepped rising pitch accent L+¡H* (for echo questions). Recent results from a Mismatch Negativity (MMN) analysis by Borràs-Comes et al. [15] using this same Catalan data back up this analysis. This study found a stronger MMN brain response when contrasting contours were presented (see Fig 3) than when listeners heard pairs of more distant non-contrasting contours.

On the other hand, the results from the Reaction Time measurements clearly show that the difference between the **statement interpretation** and the **contrastive focus interpretation** is more gradient-like. In our study, the absence of RT peaks suggests the nonexistence of categorical perception in this particular contrast. That means that the use of pitch scaling differences are not crucial for identifying narrow contrastive focus. This result is in accordance with Vanrell et al.'s [16] recent production study comparing the tonal realization of narrow contrastive focus and broad focus declaratives in Catalan and Spanish. The study shows that Catalan and two Spanish narrow contrastive focus pitch accents have (a) earlier f0 peaks and (b) higher f0 peaks than broad focus, but crucially they only have higher f0 peaks for some speakers. Further perception experiments should investigate a potential finer distinction in pitch range that might distinguish contrastive focus from echo question interpretations and also explore other cues that might better differentiate between one interpretation and the other (e.g. duration, facial gestures, presence of post-focal elements, etc.).

Methodologically, our results back up the claim by [12] and [13] that in order to interpret the results of binary identification tasks it is crucial to analyze RT patterns. According to Chen [12], "by combining the response frequencies with the mean RTs we can distinguish the task-induced identification categories from linguistically real identification categories". The absence of RT peaks in our results points to the nonexistence of categorical perception in this particular contrast.

Finally, our study highlights the fact that in Catalan increasing the pitch height of a L+H* pitch accent can have both a categorical effect (by signalling the phonological difference between a statement and an echo question), and a gradient effect (by signalling the difference between a statement and a contrastive focus interpretation).

5. Acknowledgements

We are grateful to the members of the *Group of Prosodic Studies* (UPF-UAB, Barcelona) for their comments on an earlier version of this manuscript. This research has been funded by projects FFI2009-07648/FILO and CONSOLIDER-INGENIO 2010 CSD2007-00012" (awarded by the Ministerio de Ciencia e Innovación) and by project 2009 SGR 701 (awarded by the Generalitat de Catalunya).

6. References

[1] Prieto, P; Cabré, T. [coords.] (2008). *Atles interactiu de l'entonació del català*. On line <http://prosodia.upf.edu/atlesentonacio/>

[2] Prieto, P., "Entonació" in J. Solà, M.R. Lloret, J. Mascaró and M. Pérez Saldanya, [Eds], *Gramàtica del català contemporani*, Barcelona, Editorial Empúries, 393-462, 2002.

[3] Ladd, R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.

[4] Gussenhoven, C. (1999). Discreteness and gradience in intonational contrasts. *Language and Speech* 42 (2-3), 283-305.

[5] Ladd, R.; Morton, R. (1997). The perception of intonational emphasis: continuous or categorical? *Journal of Phonetics* 25, 313-342.

[6] Francis, A. L.; Ciocca, V.; Kei Chit Ng, B. (2003). On the (non)categorical perception of lexical tones. *Perception & Psychophysics* 65 (7), 1029-1044.

[7] Savino, M.; Grice, M. (2007). The role of pitch range in realising pragmatic contrasts – The case of two question types in Italian. *Proceeding of ICPhS XVI*. Saarbrücken.

[8] Vanrell, M. M. 2006. The phonological role of tonal scaling in Majorcan Catalan interrogatives, MA thesis. UAB, Barcelona.

[9] Vanrell, M. M.; Mascaró, I.; Prieto, P., Torres-Tamarit, F. (2009). The role of pitch height in constraining the inferential space in Catalan yes-no questions. *Workshop on Prosody and Meaning*. Barcelona, 2009.

[10] Boersma, P., Weenink, D. (2008). *Praat: doing phonetics by computer* (version 5.0.09). Computer Program. On line < http://www.fon.hum.uva.nl/praat/>

[11] Psychology Software Tools. (2009). *E-Prime* (version 2.0). Computer Program. On line < http://www.pst-net.com/>

[12] Chen, A., "Reaction Time as an Indicator of Discrete Intonational Contrasts in English", EuroSpeech Proc., 97-100, 2003.

[13] Pisoni, D. B.; Tash, J. (1974). Reaction times to comparisons within and across phonetic categories. *Perception & Psycholinguistics* 15 (2), 284-290.

[14] Falé, I.; Hub Faria, I. (2005). Intonational contrasts in EP: a categorical perception approach. *INTERSPEECH-2005*, 1705-1708.

[15] Borràs-Comes, J.; Costa-Faidella, J.; Prieto, P.; Escera, C. (2009). Encoding of intonational contrasts as revealed with the mismatch negativity (MMN). *II BRAINGLOT Workshop*. Barcelona, 2009.

[16] Vanrell, M.M.; Stella, A; Gili-Fivela, B.; Prieto, P. (2010). The contribution of alignment and scaling to the perception of contrastive focus in Catalan, Italian and Spanish. Submitted to the Colloquium of Generative Grammar, Barcelona 2010.