When intonation plays the main character: information- vs. confirmation-seeking questions in Majorcan Catalan

Maria del Mar Vanrell[1,2], *Ignasi Mascaró*[3] *Francesc Torres-Tamarit*[2] *and Pilar Prieto*[1,4]

[1]Departament de Traducció i Ciències del Llenguatge, Universitat Pompeu Fabra, Barcelona, Spain
[2]Centre de Lingüística Teòrica, Universitat Autònoma de Barcelona, Bellaterra, Spain
[3]Secció de Llengua i Literatura, Institut Menorquí d'Estudis, Ciutadella, Spain
[4] ICREA, Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain

mariadelmar.vanrell@upf.edu, ipmascaro@gmail.com, francescjosep.torres@uab.cat, pilar.prieto@upf.edu

Abstract

This paper explores whether information- and confirmation-seeking questions are marked intonationally in Majorcan Catalan by means of three different perception tests (a congruity test, a rating test and a test based on the classical CP paradigm). The results show that a difference in pitch scaling on the leading H tone of the H+L* nuclear pitch accent is the main cue used by Majorcan listeners to identify confirmatory questions. Thus, while a ¡H+L* pitch accent signals an information-seeking question (i.e., the speaker has no expectation about the nature of the answer), H+L* pitch accent indicates that the speaker is asking about mutually shared information.

Index Terms: information-seeking questions, confirmation-seeking questions, tonal perception, tonal scaling, Catalan language.

1. Introduction

As is well-known, there is a well-established distinction between **information-seeking questions** (or questions that ask for information that is not mutually shared) and **confirmation-seeking questions** (or questions that ask for mutually shared information; see [1]). Thus, in confirmation-seeking questions the speaker has an expectation of the answer and is consequently asking for information that is, in a way, given or presupposed. Studies that have applied the Map Task technique ([2]) for collecting interrogative data have referred to information- and confirmation-seeking questions as queries and checks respectively ([3], [4], and [5]). In English, queries and checks can display different syntactic patterns. While canonical information-seeking questions are characterized by subject-verb inversion and the presence of an auxiliary item (*Did Jim leave early?*), canonical confirmation-seeking questions tend to display declarative syntax (*Jim left early?*).

Since in Catalan and other Romance languages these two questions are not syntactically distinct, the decision about the informational status of a question may rely heavily upon prosodic and intonational features. For example, [5] demonstrated that the degree of confidence with which the speaker believes the information to be shared with the interlocutor is reflected in Bari Italian in the choice of a specific pitch accent. In this specific case, analyses of the Map Task recordings of Bari Italian showed that questions about new information were marked by means of a rising pitch accent while questions about given information were signaled through falling pitch accents. For European Portuguese, [6] concluded that confirmation-seeking questions are not associated with a single intonation pattern. Rather, pragmatic subtypes of confirmation-seeking questions affect pitch accent types, such as questions asking for confirmation of understanding are characterized by a L+H* nuclear accent while questions requiring confirmation of perception use a H*/L+H* nuclear accent.

Regarding Eastern Catalan (Central and Balearic Catalan), a production experiment was conducted in which information- and confirmation-seeking questions were elicited by means of an intonation survey ([7]). This intonation survey was based on that developed by [8] and was designed to evoke everyday situations. It is an inductive method in which the researcher presents the subject with a series of situations. An example of the situation intended to elicit each question type follows:
-Information-seeking question. "You have a sore throat. Ask your friend whether he has a cough drop". Target question: *Do you have a cough drop?*
-Confirmation-seeking question. "A friend of yours has bought cough drops for you because you had requested it. Ask your friend whether he's bringing the cough drops". Target question: *You're bringing the cough drops?*

A total of 120 contours were obtained (3 situations x 2 question types x 20 speakers). The results revealed that there are different strategies employed depending on the dialect. While Central Catalan and Ibizan/Formenteran Catalan use boundary tones to mark the relevant distinction, Majorcan and Minorcan Catalan use different types of nuclear accents.

2. Aims

This empirical investigation focuses on Majorcan Catalan, a Catalan dialect in which the difference between information- and confirmation-seeking questions is marked through the use of the following two nuclear pitch accents respectively (see Figure 1).

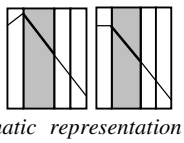


Figure 1: *Schematic representation of the nuclear accents found in Majorcan information- (left) and confirmation-seeking questions (right) (shaded area represents the stressed syllable).*

The first goal of the experiments presented here was to investigate whether the height of the leading tone in the H+L*

nuclear accent is the main perceptual cue to the contrast between information- vs. confirmation-seeking questions. The second goal was to shed light on the nature of this contrast by investigating whether this difference is categorical or gradual. To this end, three different perception experiments were conducted: a congruity test, a rating test and a test based on the application of the classical Categorical Perception paradigm.

3. Methodology

3.1. General Methodology

Twenty native speakers of Majorcan Catalan aged between 16 and 35 participated in this experiment. None of them reported a history of hearing disability. Subjects were seated at a laptop in a quiet room and the stimuli were played back through headphones. The perception tests were played by means of E-Prime and lasted a total of approximately 50 minutes. Since we were interested in Reaction Time measurements, listeners were instructed to maintain their hands near the keyboard and to press the keys as fast as they could.

3.2. Congruity Test Methodology

A congruity test was first conducted to evaluate whether the appropriate nuclear configurations could be detected in two different pragmatic contexts ([9]). Two similar everyday contexts were used for establishing the contextual appropriateness of the confirmatory/non-confirmatory meanings. In both contexts listeners had to imagine that they had just entered a store and wanted to ask the shopkeeper whether he had tangerines. They then heard on audio recordings of the ensuing dialog. The two variations of the dialog were as follows:

(1) *Només has mester mandarines, però no saps si en tenen.*
'You only need tangerines but you do not know whether they have them or not '
—*Bon dia, teniu mandarines?*
—'Good morning. Do you have tangerines?'
—*Eh… sí, ara vénc des Mercapalma i n'he duites.*
—'Er… yeah, I've just come from the wholesale market and have brought some.'

(2) *Només has mester mandarines i saps que sempre en tenen.*
'You only need tangerines and you know that they always have them.'
—*Bon dia, teniu mandarines?*
—'Good morning. You have tangerines?'
—*Clar! Com sempre!*
—'Of course, as always!'

In half of the recorded dialogs, question types were consistent with the pragmatic context. In the other half, the information-seeking question intonation was inserted into the confirmation-seeking context or vice versa. Thus, the test consisted of two yes-no questions whose intonation was coherent with the pragmatic context and two yes-no questions whose intonation was not coherent with the pragmatic context. Listeners had to answer whether they regarded the intonation of the yes-no questions as "congruent" (by pressing the "C" key) or "incongruent" (by pressing the "I" key) with the pragmatic context. The test consisted of 40 trials (2 congruous/incongruous contexts x 5 repetitions x 2 blocks).

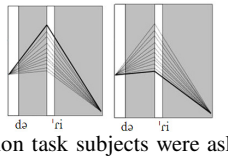
3.3. Rating Test Methodology

In the rating test, listeners had to rate on a 4 point scale the degree of presupposition that the speaker in an unmanipulated audio recording revealed about the likelihood that s/he would get a "yes" answer to her/his utterance. Subjects had four options, namely, "1" for 'no idea', "2" for 'maybe yes', "3" for 'probably yes' and "4" for 'absolutely yes', with number values reflecting the strength of certainty of a "yes" answer. The material consisted of 4 different audio stimuli: an information-seeking question (*Teniu mandarines?* 'Do you have tangerines?'), a confirmation-seeking question (*Teniu mandarines?* 'You have tangerines?', with the appropriate intonation contour), a tag question (*Teniu mandarines, no?* 'You have tangerines, don't you?'), and a declarative (*Teniu mandarines* 'You have tangerines.'). The test consisted of 40 trials (4 stimuli x 5 repetitions x 2 blocks).

3.4. Classical CP Test Methodology

Two sets of stimuli were created from two original tokens, e.g., one token of the information-seeking question *Teniu mandarines?* ('Do you have tangerines?') and one token of the confirmation-seeking question *Teniu mandarines?* ('You have tangerines?'). To create the stimuli, the f0 value at the end of the prenuclear syllable was manipulated using the resynthesis script in Praat ([10]). For each set, eleven stimuli were created by shifting the peak downwards from the information-seeking question token in 11 steps of 11.2 Hz each, and conversely by shifting the peak upwards from the confirmation-seeking question token. Differences such as the initial pitch height and the pitch height of the prenuclear accent between the two natural tokens were neutralized in order to control for all the differences not related to the height of the nuclear region. The two graphs in Figure 2 schematically show the stimuli that were created. These stimuli were used for the identification and discrimination tasks.

Figure 2: *Schematized creation of the stimuli from information-seeking question (left panel) and confirmation-seeking question base stimulus (right panel).*



In the identification task subjects were asked to respond after each stimulus by indicating the answer the speaker expected as revealed by his/her intonation. Thus, they had to press the "N" key for 'no idea' or the "M" key for 'maybe yes'. The materials for the identification task consisted of 110 trials (11 stimuli x 2 base tokens x 5 blocks).

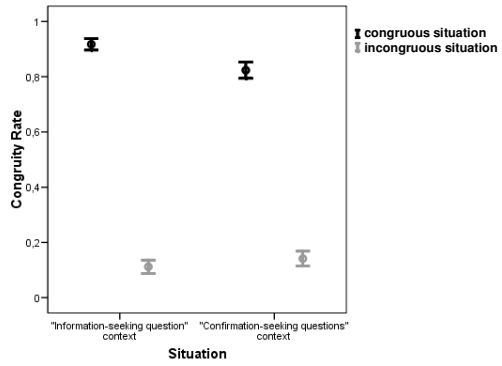
The materials for the discrimination task consisted of pairs of the same stimuli that were used in the identification task. First 20 pairs of stimuli were created in low-high order, meaning that the peak of the second stimulus is always higher in pitch than that of the first stimulus (10 from the information-seeking question and 10 from the confirmation seeking-question). Then 20 high-low-ordered pairs of stimuli were created (again, 10 for each type of question). Finally, 22 pairs which contained identical stimuli were created. As they performed the task, listeners were asked to decide whether they heard each recorded pair of stimuli as "same" or "different". The discrimination test consisted of 248 trials (10 low-high pairs + 10 high-low pairs + 11 identical stimuli pairs x 2 base stimuli x 4 blocks).

4. Results

4.1. Congruity Test Results

Figure 3 shows the rate of "congruous" responses to both congruous (black bars) and incongruous contexts (grey bars), separated into information-seeking (left) and confirmation-seeking meanings (right). The results reveal an average rate of 0.92 and 0.82 for "congruous" responses to congruous contexts in information-seeking and confirmation-seeking meanings respectively. By contrast, the average rate of "congruous" responses in the incongruous context was 0.11 and 0.14 for the information-seeking and confirmation-seeking meanings respectively. Results from the Wilcoxon matched pairs signed rank test revealed that the differences between the two conditions (congruous versus incongruous) for both information- (T=1780, p< .001, r=-.57) and confirmation-seeking question (T= 1053, p < .001, r= -.55) meanings were significant. These results indicate that listeners are extremely sensitive to the incongruous use of confirmatory and information-seeking questions.

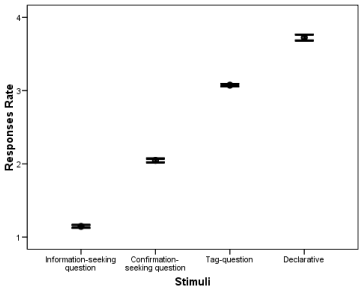
Figure 3: *Rate of "congruous" responses to congruous (black bars) and incongruous situations (grey bars), separated by information-seeking question meaning (left) and confirmation-seeking question meaning (right).*



4.2. Rating Test Results

Figure 4 shows the average "presupposition of a yes answer" score for each stimulus in the rating test with results for information-seeking question, the confirmation-seeking question, the tag question and the declarative (shown from left to right along the horizontal axis). The results show an average presupposition score of 1.15 for information-seeking questions, 2.05 for confirmation-seeking questions, 3.08 for tag questions and 3.72 for declaratives. Results from the Friedman test revealed that the differences between the scores obtained for each stimulus were significant (($\chi^2(6)$= 984.482, p<.05). This provides clear evidence that listeners base their decisions about the truth value of the sentences on not only morphosyntactic but also prosodic cues.

Figure 4: *Average presupposition score for each stimulus.*

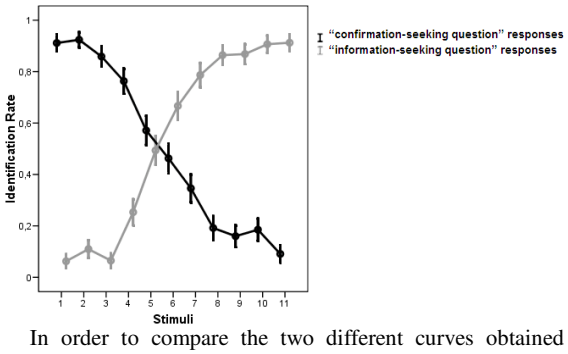


4.3. Classical CP Test Results

4.3.1. Identification Results

Figure 5 shows the identification rate for the continuum created from the confirmation-seeking question base token (black bars) and the information-seeking question base token (grey bars). The "identification rate" is defined as the number of "confirmation-seeking question" responses (in confirmation-seeking-question-based stimuli) or "information-seeking question" responses (in information-seeking-question-based stimuli). Figure 5 shows that the functions present the expected S-shape.

Figure 5: *Identification rate for the continuum created from the confirmation-seeking question (black bars) and information-seeking question base tokens (grey bars).*

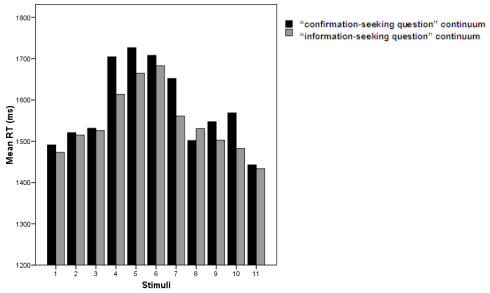


In order to compare the two different curves obtained respectively for the confirmation- and information-seeking question continua, the set of data points was fitted with a logistic function in SPSS program. From the SPSS fitted logistic curves, we obtained the boundary values calculated from the "b0" and "b1" values given for the logistic curve using the following formula: boundary= -ln(b0)/ln(b1). For the confirmation-seeking-question-based continuum, when "y" equals 0.5 "x" is 5.31, and for the information-seeking-question-based continuum, when "y" equals 0.5 "x" is 5.8. Thus the boundary is located between stimuli 5 and 6 for both continua (see Figure 5).

4.3.2. Reaction Time Results

Reaction Time (RT) measurements have been proposed to be a good alternative to the discrimination task in testing the hypothetical discreetness of a contrast ([11], [12]). [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". Figure 6 plots averaged RT responses (in ms) for all subjects. Black bars show the reaction times for the confirmation-seeking-question-based continuum and grey bars show the reaction times for the information-seeking-question-based continuum.

Figure 6: *Averaged reaction time (RT) responses (in ms) for all subjects.*

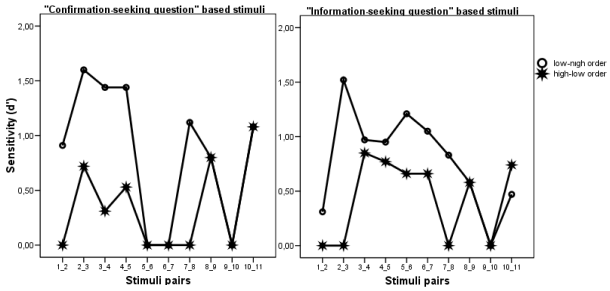


A clear peak in RT measurements for the information-seeking question continuum is obtained at stimulus 6, but not for the confirmation-seeking question continuum, since mean RT measurements for stimuli 4-6 are balanced. The location of these peaks coincides with the boundaries calculated from the fitted logistic curves. A Friedman test analysis of the reaction time results points to significant statistical differences between stimuli for both the confirmation- ($\chi^2(10)= 527.578$, $p<.001$) and information-seeking ($\chi^2(10)= 461.008$, $p<.001$) question-based continuum.

4.3.3. Discrimination Results

Figure 7 shows the discrimination results presented as d' for each stimulus pair in each order of presentation (low-high-ordered and high-low-ordered stimuli) for the confirmation (left-hand graph) and the information-seeking-question-based continua (right-hand graph). D' scores were calculated on the basis of "different" responses to the pairs that were truly different (hits) and "different" responses to the pairs that were actually the same (false alarms). Following [13], d' was calculated using roving methods (using Table A5.4, pp. 338-354). No clear peak is found at the frontier region between the categories; rather, we find two unexpected discrimination peaks at pairs 2_3 and 7_8 for the confirmation-seeking-question-based continuum and at pairs 2_3 and 5_6 for the information-seeking-question-based continuum. No match was found between this function and the identification results.

Figure 7: *Discrimination results presented as d' for each stimulus pair in each order of presentation.*



In both functions a clear order-of-presentation effect was found (low-high ordered pairs are easily perceived): a Wilcoxon matched pairs signed rank test reveals a significant difference between the two functions (low-high- ordered vs. high-low-ordered stimuli) for the confirmation (T=10829.00, $p< .001$, r=-.107) and information-seeking-question-based continua (T=1782, $p< .001$, r=-.11).

5. Conclusions

The results reported in this article confirm that a difference in pitch scaling on the leading H tone is the main cue used by Catalan listeners in distinguishing a confirmation-seeking and an information-seeking question. Thus, an upstepped leading H tone signals that the speaker has no expectation of the answer, but a non-upstepped leading H signals that the speaker is asking about mutually shared information.

Our RT measurement results indicate that pitch scaling on the H level has a phonological character. The absence of discrimination peaks should be related not to the inexistence of categorical perception but rather to a hypothetical unsuitability of the CP paradigm as applied to pitch height contrasts ([12] and [14]).

To our knowledge, this is the first perception study that proves a tonal contrast can encode knowledge presupposition in interrogative utterances. For that reason we would like to highlight the important role of perception experiments in determining the nature of the relationship between intonation and meaning.

6. Acknowledgements

Preliminary versions of this article were presented at the PaPI 2009 in Las Palmas and the Workshop on Prosody and Meaning in Barcelona. We are grateful to the participants at those events for their helpful comments and suggestions. The various Majorcans who participated unselfishly as listeners in our experiments deserve special mention. We would also thank Verònica Crespo, Marianna Nadeu and Gemma Rigau for their comments and advice on the first written version of this article. 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).

7. References

[1] Bolinger, D., "Intonation and its Uses: Melody in Grammar in Discourse", Stanford, Stanford University Press, 1989.

[2] Carletta J., Isard A., Isard S., Kowtko J. and Doherty-Sneddon G., "Corpus based approaches to dialogue modelling", Twente Workshop on Language Technology Proc., 25-34, 1995.

[3] Grice, M., Benzmueller R., Savino M and Andreeva B, "The intonation of queries and checks across languages: data from Map Task dialogues", in K. Elenius and P. Branderud [Eds], XIII International Congress of Phonetic Sciences Proc., 3:648-651, 1995.

[4] Grice M. and Savino M., "Low tone versus 'sag' in Bari Italian intonation; a perceptual experiment", XIII International Congress of Phonetic Sciences Proc., 1995.

[5] Grice M. and Savino M., "Can pitch accent type convey information-status in yes-no questions?", ACL 97 Workshop on Concept-to-Speech Generation Systems Proc., 29-38, 1997.

[6] Santos, A. L. and Mata, A.I., "Between forma and meaning: using intonation cues to identify confirmation-seeking requests", Third TIE Conference on Tone and Intonation. Universidade de Lisboa, 15-17 September, 2008.

[7] Vanrell, M.M., Mascaro, I., Prieto, P. and Torres-Tamarit, F., "Preguntar per saber i preguntar per confirmar: l'entonació de les interrogatives absolutes informatives i confirmatòries en català central i balear". Randa 64, 77-95, 2010.

[8] 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.

[9] Rathcke, T. and J. Harrington, "Is there a distinction between H+!H* and H+L* in standard German? Evidence from an acoustic and auditory analysis", Proceedings of Speech Prosody 2006, R. Hoffmann and H. Mixdorff, TUDpress Verlag der Wissenschaften GmbH, 2006.

[10] Boersma, P. and Weenink, D, Praat: doing phonetics by computer (Version 5.1.05) [Computer program]. Retrieved May 1, 2009, from http://www.praat.org/

[11] Pisoni, D. and Tash, B.J., "Same-different reaction times to consonants, vowels and syllables", Journal of the Acoustical Society of America, 55(A), 436, 1974.

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

[13] Macmillan, N.A., and Creelman, C.D., "Detection theory: a user's guide." Cambridge: Cambridge University Press, 1991.

[14] Grice, M. and Savino, M., "Reaction Time in the Perception of Intonational Contrasts in Italian", Third TIE Conference on Tone and Intonation. Universidade de Lisboa, 15-17 September, 2008.