

The role of prosody in the encoding of evidentiality

Maria del Mar Vanrell¹, Meghan Armstrong², Pilar Prieto³

¹ Institut für Romanische Philologie, Freie Universität Berlin, Berlin, Germany

² Languages, Literatures, & Cultures, University of Massachusetts-Amherst, Amherst, USA

³ Departament de Traducció i Ciències del Llenguatge, ICREA-Universitat Pompeu Fabra, Barcelona, Spain

mariadelmar.vanrell@fu-berlin.de, armstrong@spanport.umass.edu, pilar.prieto@upf.edu

Abstract

The overarching goal of this paper is to advance on the understanding of how evidential meanings are expressed in natural languages. Specifically, we aimed to investigate what type of meaning was encoded in yes-no questions through the combination of the question particle (QP) *que* ‘that’ and the nuclear intonational pattern L+H* L% in Majorcan Catalan yes-no questions (i.e., *Que és un llibre?*_{L+H* L%} ‘QP-It’s a book?’), and to understand any temporal information that might be encoded through this construction. Several complementary research methods were used to address our question: the Discourse Completion Task, an acceptability task and a multiple-choice questionnaire. The results show that three types of information are encoded in QP *que* L+H* L% questions: sentence modality, inference based on direct evidence and immediacy of the evidence.

Index Terms: intonation, question particles, evidentiality, typology, Catalan.

1. Introduction

All languages have some way of marking information source, though different parts of the grammar may be recruited to do so depending on the language. Lexical means for specifying source of information are probably universal. For instance, the lexical item *es veu* is used in Catalan in (1) to indicate that the speaker does not have firsthand knowledge of the proposition ([1]). In Colombian Spanish, the combination of the verb ‘say’¹ and the complementizer ‘that’ can have different functions such as a) reported speech, b) hearsay, c) labelling or d) dubitative ([4], [5]). This is also found in other Romance language such as Galician, Romanian, Sardinian or Sicilian ([6]).

- (1) *S’han quedat sense llum a Girona. Es veu que hi ha nevat molt.*
‘The power is out in Girona. There must have been a lot of snow.’
- (2) *Y eso, **dizque** es peligroso, no?*
‘And that’s dangerous (reported), no?’

Some languages have true evidential systems, i.e., inflectional systems with morphemes that have source-marking at the core of their semantics and are obligatory. In other languages, evidentiality specifications are “scattered”

throughout the grammar ([4]). While there is some work showing the relevance of prosodic or gestural cues in the expression of epistemicity, i.e., the degree of certainty about a proposition (see [7], [8], [9] for declaratives; [10] for both declaratives and interrogatives, and [11] for interrogatives), research examining prosodic marking of evidential strategies is still scarce. The only investigation we are aware of is [12], who showed that deaccenting is used in Japanese biased polar questions when the speaker expects a positive answer based on public evidence.

Majorcan Catalan is a variety of Catalan that is characterized by exhibiting several strategies for forming polar questions based on bias ([13], [14]). Speakers may choose from different pitch accents as in (3) or may head questions with question particles (QPs) such as *que* (complementizer ‘that’) or *o* (conjunction ‘or’), see (4).

- (3) *Teniu mandarines?*_{iH+L* L%} ‘Do you have tangerines?’ vs *Que hi ha gana?*_{L+H* L%} ‘Are you hungry?’
- (4) *Que encara no ha vengut, s’electricista?* ‘The electrician hasn’t arrived yet?’ vs *O no estàs bo?* ‘Aren’t you well?’

The main goal of this article is to assess the role intonation plays in encoding information source in questions headed by the QP *que* (i.e., in constructions of the type QP *que* L+H* L%). If it is the case that information source is being encoded through the use of QP *que* L+H* L%, we also asked what type of evidential information the speaker might be conveying. Since evidential markers may have temporal restrictions ([15]) we also sought to investigate whether, in cases where evidential information was encoded, temporal information might also be available to listeners. Some languages, such as Sherpa ([16], [17]), encode temporal information through the evidential marker. In the case of Sherpa, *immediate* evidence is encoded. We thus left this as an additional possibility, since often times questions are asked based on information that has just become available in the discourse ([18]).

We tested these questions performing different experiments: a) Experiment 1: the Discourse Completion Task ([19], [20], [21]), b) Experiment 2: the acceptability task and c) Experiment 3: the multiple-choice questionnaire.

2. Experiment 1: the Discourse Completion Task

2.1. Methodology

For this study the Discourse Completion Task involved the creation of a set of situations which contained two evidential conditions (inferred direct evidential and hearsay) plus a non-evidential situation (see examples (5), (6) and (7) respectively). The difference between the two conditions is

¹ The form *diz que* could be a shortened version of the plural form *dicen que* ‘they say that’ ([2]) which “began as the collocation of a verb introducing indirect speech and its complementizer, and developed into an evidential strategy” ([3]: 16).

that while in direct evidential contexts the speaker infers the proposition from direct evidence (s/he sees it directly), in hearsay contexts inference is based on hearsay or the report of the proposition by another individual. The situations provided in the DCT were read aloud by the first author of this paper to the participants. Then, the participants were asked to respond appropriately to them. The situations were expected to elicit two biased yes-no questions (in the case of the two evidential conditions) and a neutral polar question. The questionnaire was made up of 12 situations (3 conditions x 4 pragmatic contexts). Fifteen speakers (8 females, 7 males) participated in the experiment. The final database was comprised of 180 utterances. Data were coded in Praat ([22]) for the following: a) use of lexical markers (question particles such as *o* ‘or’, or *que*), b) use of syntactic markers (negative yes-no questions, split questions) and c) prosodic transcription using the latest version of the Cat_ToBI system ([23]). The annotations were collected automatically in .txt format through a Praat script and then transferred to an Excel and a SPSS file for subsequent statistical analysis.

- (5) Non-evidential situation: ‘You have a bit of a cough and suddenly, while you’re talking to a neighbor, you feel a sore throat coming on. Ask her if she has a cough drop’.
- (6) Hearsay situation: ‘Your cousin visits you. While you are talking, she says that since she’s stopped working, she has plenty of free time. You take this to mean that she retired, so you confirm with her’.
- (7) Direct evidential situation: ‘It’s your birthday and your friend gives you a present. You ask him if it’s a book, since the package has a book-like shape’.²

2.2. Results

Figure 1 shows the percentage of occurrence of the dependent variable COMBINATION OF LEXICAL/SYNTACTIC STRATEGIES AND INTONATION with three levels of lexical/syntactic strategies (no QP, QP and other) and four levels of intonational strategies (\uparrow H+L* L%, H+L* L%, L+H* L% and other)³ for each of the possible evidential conditions (neutral, hearsay or direct evidence). The results clearly show that whereas neutral and hearsay contexts elicit a high percentage of the \uparrow H+L* L% intonational pattern without any QP, direct evidential contexts elicit the production of the L+H* L% intonational pattern headed by a QP (*que* or *o*). A Friedman test revealed a significant effect of evidential condition (neutral, hearsay or direct evidence) on the intonation of the y/n questions ($\chi^2(2) = 28.203, p < .001$) as well as on the lexical/syntactic marking ($\chi^2(10) = 24.125, p < .001$).

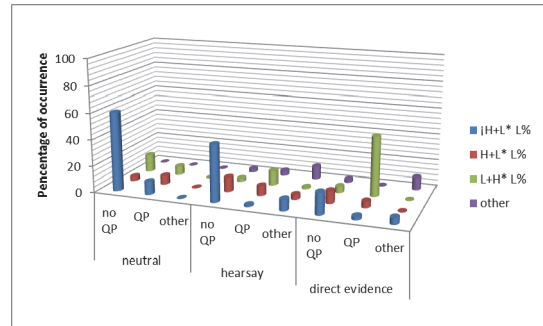


Figure 1: Percentage of occurrence of the variable COMBINATION OF LEXICAL/SYNTACTIC STRATEGIES AND INTONATION for each of the possible pragmatic contexts (neutral, hearsay or direct evidence).

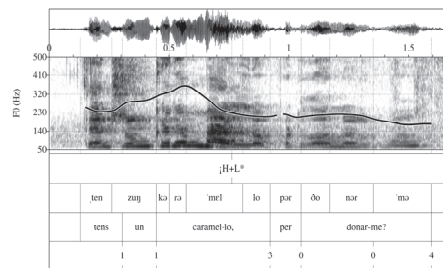


Figure 2: Waveform and F0 contour of the y/n question *Tens un caramel·lo, per donar-me?* ‘Do you have a candy to give me?’ (the most common contour in the neutral pragmatic contexts).

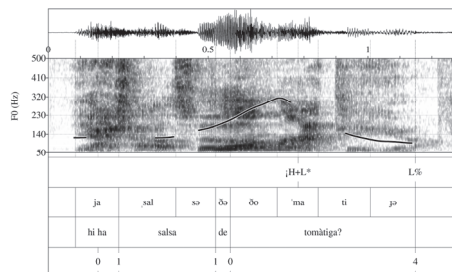
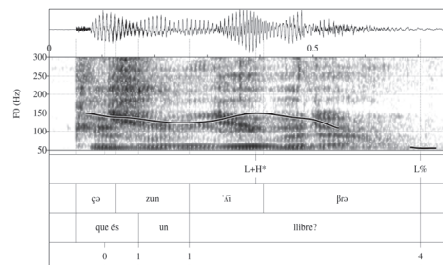


Figure 3: Waveform and F0 contour of the y/n question *Hi ha salsa de tomàtiga?* ‘There’s tomato sauce?’ (the most frequent pattern in the hearsay condition).



² All the biased situations used in the experiments had a positive bias, that is, they were situations designed to elicit an affirmative response.

³ The \uparrow H+L* L% nuclear configuration is realized as a fall with a leading extra high tone aligned with the end of the preaccentual syllable. The H+L* L% pattern is identical to the previous configuration but with a just high leading tone. The L+H* L% pattern is characterized by a rising movement with the peak aligned with the end of the accented syllable followed by a falling final movement.

Figure 4: *Waveform and F0 contour of the y/n question Que és un llibre? 'QP-It's a book?' (preferred strategy for direct evidential contexts).*

3. Experiment 2: the Acceptability task

3.1. Methodology

Experiment 2 consists of an acceptability task, which allows us to evaluate the degree of perceived appropriateness of target intonational patterns (in this specific case, with the combination of lexical markers) to different pragmatic contexts. Six pragmatic contexts were created with two different evidential conditions (direct evidential for half of the contexts and hearsay for the other half, see (8) and (9) below). Then three different combinations of intonation and QP, which were understood to be semantically appropriate given the pragmatic contexts, were produced, as follows: a) no QP_iH+L* L% intonational pattern; b) QP *que* L+H* L% intonational pattern; and c) QP *que* _iH+L* L% intonational pattern.⁴ Thus, we had a total of 36 trials: 2 contexts (direct evidential and hearsay) x 3 combinations of QP and intonational conditions (no QP_iH+L* L%, QP *que* L+H* L% and QP *que* _iH+L* L%) x 3 pragmatic contexts + 18 fillers.

Subjects were asked to rate the acceptability of the question produced in a specific context using a 7-point Likert scale (1=totally unacceptable, 7=acceptable). Forty-six Majorcan Catalan listeners (23 female and 23 male) participated in the experiment. The experiment was run by the Survey Gizmo software (online survey software – surveygizmo.com) and was approximately 10 min long.

- (8) Hearsay situation: 'Maria arrives at the fruit store and hears the owners talking about how they have the shoes from the shop owner next door, who has recently retired and closed the shop.' Target question: 'Are you going to sell shoes?'
- (9) Direct evidential situation: 'Maria arrives at the fruit store and sees that the owners are putting pairs of shoes on the shelves in the entrance.' Target question: 'Are you going to sell shoes?'

3.2. Results

Table 1 shows the mean of the ACCEPTABILITY RATINGS (dependent variable) and the standard deviation of the three possible combinations of QP and intonation (no QP_iH+L* L%, QP_iH+L* L% and QP L+H* L%) for each of the possible evidential conditions (direct evidential and hearsay). As we can observe in Table 1, listeners rate the combination of QP *que* and the L+H* L% intonational pattern as the most natural when it is produced in a direct evidential context (mean of 5.41). However, the same pattern of intonation and QP shows a high rate of acceptability in the hearsay context (mean = 5.12). The results of two Friedman tests showed a statistically significant effect of the combination of QP + intonation on the acceptability mean in hearsay ($\chi^2(2) = 15.510, p < .001$) and in direct evidence context ($\chi^2(2)$

=32.690, $p < .001$). The Wilcoxon tests using a Bonferroni correction showed that the QP *que* L+H* L% pattern differs statistically from the rest in both the direct evidential and hearsay conditions ($p < .001$)

Table 1. *Acceptability mean and ± standard deviation for each combination of QP + intonation and pragmatic context.*

	Hearsay	Direct Evidence
no QP _i H+L* L%	4.56 ±1.69	4.51 ±1.68
QP <i>que</i> _i H+L* L%	4.27 ±1.68	4.34 ±1.95
QP <i>que</i> L+H* L%	5.12 ±1.73	5.41 ±1.73

4. Experiment 3: Multiple-choice questionnaire

4.1. Methodology

For the multiple-choice questionnaire we created three pragmatically neutral contexts and three different combinations of QP and intonational conditions. Each combination of QP and intonational conditions was inserted in each of the pragmatically neutral contexts. Therefore, we had a total of 14 trials: 3 pragmatically neutral contexts x 3 QP + intonational conditions (no QP_iH+L* L%, QP *que* L+H* L% and QP *que* _iH+L* L%) + 7 fillers. The subjects were asked to answer 2 multiple-choice questions related to the information source and to the time at which the evidence was available. One example of the trials was as follows: —If Maria says 'You're going to sell shoes?' to the owners of the fruit store is because Maria: a) heard that they might sell shoes, b) saw that they might sell shoes, c) heard or saw that they might sell shoes (but I'm not sure), d) I don't know. —When did Maria hear or see it? a) Just now, b) A few hours ago, c) Yesterday, d) I don't know. The 'hearing' response was an intuitive response for the hearsay condition and the 'seeing' response for the direct evidential condition. Experimental presentation was done using the Survey Gizmo software (online survey software – surveygizmo.com) for the forty Majorcan Catalan listeners (25 females, 15 male) that participated in this task and lasted approximately 10 min.

4.2. Results

Figure 6 shows the count of the dependent variable SOURCE OF INFORMATION (four levels: auditory, visual, auditory or visual, DK) for the independent variable COMBINATION OF QP AND INTONATION (three levels: no QP_iH+L* L%, QP *que* _iH+L* L% and QP *que* L+H* L%). The results show that the QP *que* L+H* L% condition elicits the highest rate of visual (54 out of 360 responses) and visual/auditory (48 out of 360) responses and no "I don't know" responses. By contrast, the QP *que* _iH+L* L% condition elicits the highest rate of auditory responses (43 out of 360) and the second highest rate of visual/auditory responses (41 out of 360). The effect of the combination of QP and intonation was not statistically significant (Friedman test: $\chi^2(2) = 4.516, p > .05$). Wilcoxon tests were used to follow up on this finding and they demonstrated that the effect of the QP *que* L+H* L% significantly differed from the effect of the _iH+L* L% intonational pattern ($T = 1019.50, p < .05, r = -.22.89$) on the count of the dependent variable SOURCE OF INFORMATION.

⁴ The reason why we did not fully cross the lexical and intonational conditions is because the combination of the QP *que* and the L+H* L% pattern is ungrammatical. Thus, we prevented the listeners from rating the degree of acceptability of the target itself, rather than that of the target produced in a specific pragmatic context.

Figure 7 shows the count for the dependent variable TIME OF EVIDENCE (four levels: just now, a few hours ago, yesterday, DK) for the independent variable COMBINATION OF QP AND INTONATION (three levels: no QP_iH+L* L%, QP *que*_iH+L* L% and QP *que*_LH+L* L%). As observed in the graph, the highest rate of 'right now' responses (70/360) were obtained for the condition QP *que*_LH+L* L%. The Friedman analysis showed that the combination of lexical and intonational strategies had a statistically significant effect on the dependent variable TIME OF EVIDENCE (Friedman test: $\chi^2(2) = 13.225, p = .001$). The subsequent Wilcoxon tests revealed statistically significant differences between the combination of QP_LH+L* L% and the rest of conditions ($p < .01$).

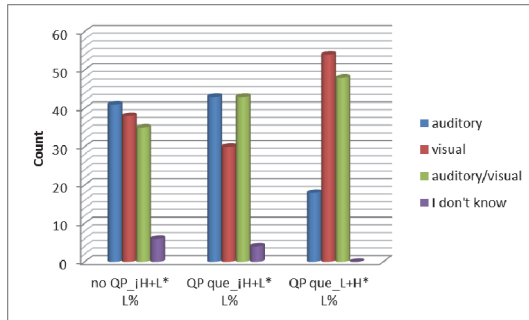


Figure 6: Count of the dependent variable SOURCE OF INFORMATION for each combination of QP + intonation.

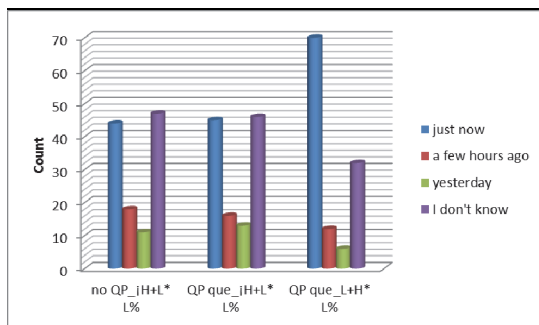


Figure 7: Count of the dependent variable TIME OF EVIDENCE for each combination of QP + intonation.

5. Discussion and conclusions

The aim of this study was to investigate what type of meaning was encoded through the combination of the QP *que* and the nuclear pattern L+H* L% in Majorcan Catalan y/n questions, and also to understand any related temporal information being conveyed through this construction. To this end, a set of three experimental tasks were carried out: a production experiment using the Discourse Completion Task methodology, a perception test (acceptability task) and another perception test based on a multiple-choice questionnaire. The results obtained from these different experimental approaches allow us to conclude that three types of information are encoded in QP *que*_LH+L* L% questions. First, sentence modality, specifically interrogative modality, is encoded by the combination of the question particle and the rising-falling intonational pattern. Second, this particular construction informs us that the speaker has inferred the proposition through direct evidence (visual, in

the specific case of this work) and now wants to ask for confirmation of this inference. Finally, listeners that hear the QP *que*_LH+L* L% construction infer that the propositional content became available to the speaker just prior to the time of the utterance. Therefore, listeners are able to infer information about time reference directly through the QP *que*_LH+L* L%. and do not need to rely on pragmatic context for this information. The construction we describe, then, also conveys immediate evidence like the Sherpa case mentioned above.

In this paper, only visual direct information was tested for direct evidential contexts. For instance, [15] showed that sentences with the particle *-te* in Korean signal sensory evidence, regardless of the tense that they occur with. We suspect that this is also the case for QP *que*_LH+L* L% questions, in Majorcan Catalan. Unlike the *-te* particle, however, our data show a temporal restriction such that this construction can only be used just after the proposition was activated, and as a result listeners infer this temporal information when they hear the construction. Further research will be necessary to test whether the results obtained for the sense of sight are also applicable to other senses.

In sum, we conclude that the QP *que*_LH+L* L% questions work as a construction, that is, a learned pairing of form and meaning ([24], [25]) which conveys sentence modality, source-marking and temporal information. In her paper on Korean evidentials, [15] notes the various strategies for encoding evidential information reported by [26] in his survey of 418 languages: verbal affix/clitic (131 languages) or particles (65 languages). [15] calls for research on languages that do not express evidentiality through distinct evidential morphemes, proposing that this will help us to understand how evidential meanings are expressed and how they make specific evidential distinctions. Our results show that without recruiting a specific part of the grammar (intonation), source-marking and temporal information are no longer available to the listener. Thus, we confirm that intonation is indeed a part of the grammar available for conveying evidential meaning, in this case working in tandem with the particle *que*. Our findings therefore have important typological implications for the study of evidential meaning.

6. Acknowledgments

A preliminary version of this paper was presented at the conference *Phonetics and Phonology in Iberia* (June 2013, Lisboa, Portugal) and the conference *Hispanic Linguistics Symposium* (October 2013, Ottawa, Canada). We are grateful to the participants at these conferences for their helpful comments and suggestions. We also thank all the subjects that participated in the different experiments. This research was funded by projects FFI2011-23829/FILO, BFI2012-31995 (awarded by the Spanish Ministry of Science and Innovation and the Spanish Ministry of Economy and Competitiveness) and 2009 SGR 701 (awarded by the Generalitat de Catalunya).

7. References

- [1] González, M., “Indirect evidence in Catalan: A case study”, in Ll. Payrató and Cots, J.M. [Eds.], *The Pragmatics of Catalan*, 146-172, Mouton de Gruyter, 2012.
- [2] López-Izquierdo, M., “L’emergence de dizque comme stratégie médiative en espagnol médiéval”, *Cahiers d’études hispaniques médiévales*, 29: 483-493, 2006.
- [3] Miglio, V.G., “Online databases and language change: the case of Spanish dizque”, St.Th. Gries, St. Wulff and Davies, M. [Eds.], *Corpus-linguistic Applications. Current studies, new directions*, 7-28, Rodopi, 2009.
- [4] Aikhenvald, A.Y., *Evidentiality*, Oxford University Press, 2004.
- [5] Travis, C., “Dizque: a Colombian evidentiality strategy”, *Linguistics*, 44(6): 1269-1297, 2006.
- [6] Cruschina, S. and Remberger, E., “Hearsay and reported speech: Evidentiality in Romance”, *Rivista di Grammatica Generativa*, 33: 95-116, 2008.
- [7] Swerts, M. and Krahmer, E., “Audiovisual prosody and feeling of knowing”, *Journal of Memory and Language*, 53: 81-94, 2005.
- [8] Dijkstra, C., Krahmer, E. and Swerts, M., “Manipulating uncertainty: The contribution of different audiovisual prosodic cues to the perception of confidence, in R. Hoffmann and Mixdorff, H. [Eds.], *Proceedings of the Third International Conference on Speech Prosody*, 1-4, Dresden, 2006.
- [9] Borràs-Comes, J., Roseano, P., Vanrell, M.M., Chen, A. and Prieto, P., “Perceiving uncertainty: facial gestures, intonation, and lexical choice”, in C. Kirchhof, Z. Malisz and Wagner, P. [Eds.], *Proceedings of the 2n Conference on Gesture and Speech in Interaction*, 2011.
- [10] Gravano, A., Benus, S., Hirschberg, J., German, E.S. and Ward, G., “The effect of prosody and semantic modality on the assessment of speaker certainty”, in P. Barbosa, S. Madureira and Reiss, C. [Eds.], *Proceedings of the fourth Speech Prosody 2008 Conference*, 401-404, 2008.
- [11] Vanrell, M.M., Mascaró, I., Torres-Tamarit, F. and Prieto, P., *Intonation as an Encoder of Speaker Certainty: Information and Confirmation Yes-No Questions in Catalan*, *Language and Speech*, 56(2), 163-190, 2013.
- [12] Hara, Y., Kawahara, S., “The prosody of public evidence in Japanese: A rating study, in Ch. Jaehoon, A. Hogue, J. Punske, D. Tat, J. Schertz and Trueman, A. [Eds.], *Proceedings of the 29th West Coast Conference on Formal Linguistics*, 353-361, Cascadilla Press, 2012.
- [13] Prieto, P. and Cabré, C. [Ed.], *Atles interactiu de l’entonació del català*. Online: <http://prosodia.upf.edu/atlesentonacio/>, accessed on October 2013, 2007-2012.
- [14] Vanrell, M.M. and Mascaró, I., “Balear”, in Prieto, P. and Cabré, T. [Ed.], *L’entonació dels dialectes catalans*, 75-100, Publicacions de l’Abadia de Montserrat, 2013.
- [15] Lee, J., “Temporal constraints on the meaning of evidentiality”, *Nat Lang Semantics*, 21: 1-41, 2013.
- [16] Woodbury, T., “Interactions of tense and evidentiality: A study of Sherpa and English, in W. Chafe and Nichols, J., *Evidentiality: The linguistic encoding of epistemology*, 188-202, Ablex, 1986
- [17] Kelly, B., *A grammar and glossary of the Sherpa language*, in *Tibeto-Burman languages of Nepal*: Manange and Sherpa, 197-324, The Australian National University, 2004.
- [18] Büring, D. and Gunlogson, C. “Aren’t positive and negative polar questions the same?”, paper presented at the LSA Annual Meeting, Chicago, Ms., University of California, Los Angeles and University of California, Santa Cruz. Online: <http://hdl.handle.net/1802/1432>, accessed on January 2011, 2000.
- [19] Blum-Kulka, S., House, J. and Kasper, G., “Investigating cross-cultural pragmatics: An introductory overview”, in S. Blum-Kulka, J. House and Kasper, G. [Eds.], *Cross-cultural pragmatics: Requests and apologies*, 13-14, Norwood, NJ: Ablex, 1989.
- [20] Billmyer, K. and Varghese, M., “Investigating instrument-based pragmatic variability: Effects of enhancing discourse completion tests”, *Applied Linguistics*, 21(4):517-552, 2000.
- [21] Félix-Brasdefer, C., “Data collection methods in speech act performance: DCTs, role plays, and verbal reports”, in A. Martínez-Flor and Usó-Juan, E. [Eds.], *Speech act performance: Theoretical, empirical, and methodological issues*, 41-56, John Benjamins, 2010.
- [22] Boersma, P. and Weenink, D., *Praat: Doing phonetics by computer [Computer program]*, Version 5.3.55. Online: <http://www.praat.org>, accessed on 2 Sep 2013.
- [23] Prieto, P., Borràs-Comes, J., Cabré, T., Crespo-Sendra, V., Mascaró, I., Roseano, P., Sichel-Bazin, R. and Vanrell, M.M., “Intonational phonology of Catalan and its dialectal varieties”, in S. Frota and Prieto, P. [Eds.], *Intonational variation in Romance*, OUP, in press, to appear in 2014.
- [24] Goldberg, A., *Constructions*, University of Chicago Press, 1995.
- [25] Goldberg, A., *Constructions at work*, Cambridge University Press, 2006.
- [26] De Haan, F., “Coding of evidentiality”, in M. Haspelmath, M. Dryer, D. Gil and Comrie, B. [Eds.], *The world atlas of language structures*, 318-323, Oxford University Press, 2005.