

On the Intonation of Confirmation-Seeking Requests in Child-Directed Speech

Ana Isabel Mata, Ana Lúcia Santos

FLUL / CLUL, University of Lisbon, Lisbon, Portugal

{aim, als}@fl.ul.pt

Abstract

In this paper we identify intonation cues that can disambiguate confirmation-seeking questions in adult-child dialogue in European Portuguese (EP). 301 examples of confirmation requests answered by two children and uttered by three different adults were analysed. Results show that (i) most confirmation-seeking questions (92.7%) do not present the intonation pattern previously identified for information-seeking questions in EP; (ii) pragmatic/discourse values of confirmation-seeking questions affect pitch accent type distribution and F0 height of both nuclear pitch accents and final boundary tones; (iii) L*+H and ^H*, previously associated with narrow/contrastive focus in questions or with correction of given information, are associated with non-neutral acceptance in confirmation requests. We interpret non-neutral acceptance as an instance of Contrast and suggest that Contrast is coded across different contexts and structures by the same pitch accents.

Index Terms: intonation, confirmation-seeking questions, child-directed speech

1. Introduction

Yes-no questions seeking confirmation, as exemplified by A's request in (1), may have different interpretations: the speaker may be asking *about* the little duck (seeking confirmation of understanding), or he may be asking *about* what was said, trying to confirm what the speaker actually uttered (seeking confirmation of perception).

- (1) C(hild): O patinho tem um amigo.
the little duck has a friend
'The little duck has a friend.'
A(dult): O patinho tem um amigo?
the little duck has a friend
'Does the little duck have a friend?'
C(hild): a. Sim. / b. É. / c. Tem.
yes / is (=yes) / has (=yes)

These different interpretations may correlate with different affirmative answer patterns. In the case of a perception reading, the affirmative answer corresponding to a stranded finite verb in a VP ellipsis or null object structure (*tem* 'has') is excluded: this answer makes an irrelevant assertion given the interpretation of the question. In the case of an understanding reading, all the answer types are acceptable [1].

These yes-no questions, which may function as a confirmation-seeking request (not an information-seeking one), may be identified by the discourse context. Are they also identified by intonation cues in EP? And, more specifically, do different interpretations of confirmation-seeking questions (perception / understanding) correlate with different intonation patterns?

Based on a Map-Task corpus, [2] suggested that intonation in EP may distinguish information-seeking from confirmation-

seeking questions. However, there is no work on EP proving an association between intonation patterns and pragmatic/discourse subtypes of confirmation-seeking questions (e.g. confirmation of perception vs. confirmation of understanding).

Recent research on languages such as English, German, Italian or Swedish has shown that intonation can distinguish questions according to different discourse / pragmatic values (function):

- **Confirmation-seeking and information-seeking questions** – Pragmatically different subtypes of confirmation requests affect the distribution of rising / falling contour or pitch accent types ([3] for Italian; [4] for German, Italian and Bulgarian).
- **Confirmation, clarification and information seeking questions** - Type of clarification sought affects the distribution of rising / falling boundary tones ([5] for German); high peak location disambiguates between questions centered on perception and those centered on understanding ([6] for Swedish). F0 height (nuclear accent and boundary tone) varies according to the function of questions (information > clarification > confirmation) and according to the type of clarification (acoustic/ perceptual > understanding >...> intention) ([7] for English).

In this paper, we provide evidence that (i) EP confirmation-seeking questions and information-seeking questions exhibit predominantly different intonation patterns; (ii) the mechanisms disambiguating the interpretation of different subtypes of confirmation questions may include type of pitch accent and F0 height – confirmation of perception vs. confirmation of understanding; in addition, (iii) the distribution of pitch accents in subtypes of understanding confirmation-seeking questions is a cue to different discourse/pragmatic values, which are determined by levels of acceptance / contrast.

2. Methodology

2.1. Corpus selection

We selected and coded a corpus of child-directed speech collected by [1], which by its nature presents a broad range of communication problems favouring the occurrence of a large variety of confirmation requests. We analysed a subset of 23h of spontaneous dialogue between adults and two children and selected 307 confirmation-seeking questions answered by children and uttered by three different adults. 6 out of the 307 questions were wh-questions which call for yes-no answers and are consequently interpreted as yes-no questions (asking for confirmation of perception); given their low occurrence, these wh- confirmation-seeking questions are not included in the present study. The questions analysed include 143 sentences (mean number of words: 3; mean number of syllables: 4.8) and 158 fragments (mean number of words: 1.8;

mean number of syllables: 2.8), fragments being verbless utterances as opposed to sentences. Sentential yes-no questions present declarative word order in Portuguese.

2.2. Coding of confirmation-seeking questions

Pragmatic / discourse coding: Previous literature is not consensual in the definition of questions' pragmatic functions (see discussion in [8]). Given that the definition of confirmation requests and, particularly, the distinction between confirmation and clarification is not always clear, we limited our research to cases that may be defined as confirmation requests according to [9], who distinguishes clarification and confirmation. [9] assumes that clarification requests call for more precise information or for justification (as in 2); confirmation requests signal that the speaker makes an assumption about the answer to the question but he wants it to be confirmed (as in 3).

(2) A: You should take Lisbon street.

B: Lisbon street? [meaning: I don't know this street, can you explain me better where it is?]

(3) A: You should turn left.

B: Left? [meaning: did you say or meant to say "left"?]

Confirmation requests, which are the only object of this study, were tagged by one of the authors according to three main levels of action (to define the source of the problem leading to the confirmation request): understanding, perception and contact ([10]; see also the classification of the source of the problem by [5] and [7]). As stated in section 1, if the confirmation request is related to 'understanding', the question is related to the meaning of what was previously uttered; if the confirmation request is related to 'perception', it is an attempt to confirm what was pronounced. On the contrary, contact questions do not result from a real communication problem: the speaker asks for a confirmation only seeking to maintain the ongoing interaction.

A preliminary attempt to classify the data showed that confirmation-seeking questions related to 'understanding' should be distinguished according to level of 'acceptance', another level of action [10]. Therefore, confirmation-seeking questions related to understanding were further coded according to neutral / non-neutral acceptance (non-neutral acceptance includes the suggestion of incorrectness, disbelief or surprise). Additionally, those related to 'perception' were distinguished according to two subcategories: high/low level of confidence that the information received is correct [4].

Intonation analysis and F0 measurements: The corpus was transcribed using a system based on the first proposal *Towards a P_ToBI* [2] (see <http://www.ling.ohio-state.edu/~tobi/>). 7 of the 8 pitch accents covered in that proposal (H+L*, H*+L, L*+H, L+H*, H*, L*, ^H*) and all the final boundary tones (L%, H%, !H%, LH%, HL%) were used. F0 values corresponding to high and low targets within nuclear pitch accents and final boundary tones were measured and converted in semitones. For pitch accents, the F0 maximum (turning) point (H) and the F0 minimum (turning) point (L) were measured in or near the stressed syllable of the accented word. For boundary tones, identical measures of F0 maxima and minima were taken at the rightmost edge of the question. The difference between the maximum and minimum F0 points (F0 range) was also calculated for pitch accents and boundary tones.

The intonation coding was manually done by one of the authors and the pragmatic / discourse coding was independently done by the other.

3. Results

3.1. Nuclear pitch accents

The first type of relevant results obtained concerns the distribution of nuclear pitch accents across subtypes of confirmation-seeking questions (see Table 1).

	H*	L*+H	L+H*	H+L*	L*	^H*	H*+L	N (%)
Understanding	11	54	22	16	11	15	7	136 (45.2)
Perception	56	13	35	11	7	0	1	123 (40.9)
Contact	1	5	11	11	14	0	0	42 (14)
N (%)	68 (22.6)	72 (23.9)	68 (22.6)	38 (12.6)	32 (10.6)	15 (5)	8 (2.7)	301 (100)

Table 1. *Distribution of nuclear pitch accents across subtypes of confirmation-seeking questions*

The results allow two different types of generalization. First, even though there is a wide consensus concerning the contour of information-seeking yes-no questions in EP (H+L* LH%) (see [11] for an overview; [2] for examples from read and spontaneous speech), only 7.3% of the confirmation-seeking questions that were analysed present this pattern and only 12.6% out of 301 examples have a H+L* (the most commonly used nuclear accent in information-seeking questions in EP). This is a clear result supporting a distinction between the intonation patterns of information-seeking and confirmation-seeking yes-no questions – confirming the identification of a distinction in [2].

Second, as Table 1 shows, confirmation-seeking questions are not associated to a single intonation pattern. But crucially there is a correlation between pitch accents and the discourse/pragmatic function of this type of questions ($\chi^2(12)=134.95, p<.0005$): H* and L+H* mainly occur in confirmation-seeking questions coded as 'perception' related; L*+H mainly occurs in confirmation-seeking questions related to 'understanding'. See figures 1 and 2.

(4) (see Figure 1)

C: é o do João Abíl(io).

is the of+the João Abílio

A: o do João é o Abílio?*

the of+the João is the Abílio

C: é.

is

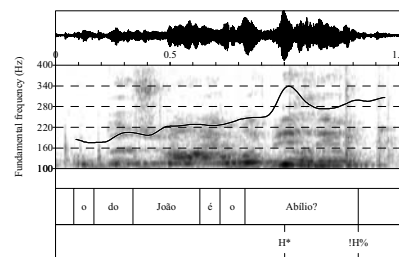


Figure 1: *H* !H% used in a confirmation-seeking question coded as 'perception'* (figure done with Praat)

* The stressed syllable of the accented word is underlined.

(5) (see Figure 2)

C: xxx # é dos # animai(s).
 is Prep+the animals
 A: dos animai[i]s? *
 Prep+the animals
 C: sim.
 yes

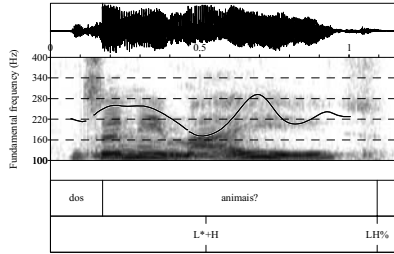


Figure 2: L*+H LH% used in a confirmation-seeking question coded as 'understanding' (figure done with Praat)

Moreover, other correlations may be identified concerning the distribution of pitch accents among pragmatic subtypes of confirmation-seeking questions coded as 'understanding' or 'perception'.

In the case of 'understanding' confirmation-seeking questions, there is a significant correlation between level of acceptance and pitch accent ($\chi^2(6)=44.31, p<.0005$). L*+H is the most frequent accent. However, when understanding questions are broken down by acceptance type, L*+H mainly occurs in questions expressing non-neutral acceptance (38 out of 54 cases, or 70.4%, of L*+H in understanding questions are cases of non-neutral acceptance). Moreover, L*+H is the most frequent pitch accent in non-neutral acceptance understanding questions (38 out of 69, or 55.1%). The other frequent pattern in non-neutral cases is ^H* (15 out of 69, or 21.7%). H+L* (the most commonly used in EP yes-no questions) mainly occurs in neutral understanding questions.

In the case of confirmation-seeking questions coded as 'perception' (Table 2), the pragmatic sub-types are also significantly correlated with pitch accents ($\chi^2(5)=24.04, p<.0005$). We notice that these are overwhelmingly marked with a H* (40 out of 67 examples, or 59.7%) when the level of confidence is high; of the 56 examples coded as presenting a low level of confidence that the information perceived is correct, 26 (46.4%) have a L+H* pitch accent and 16 (28.6%) a H* pitch accent.

	H*	L*+H	L+H*	H+L*	L*	^H*	H*+L	N (%)
Perception CL high	40	5	9	9	3	0	1	67 (54.5)
Perception CL low	16	8	26	2	4	0	0	56 (45.5)
N (%)	56 (45.5)	13 (10.6)	35 (28.5)	11 (8.9)	7 (5.7)	0 (0.8)	1 (100)	123

Table 2. Distribution of pitch accents according to levels of confidence in perception related questions

As for contact related confirmation questions, they can frequently take L* (14 / 33.3%), H+L* (11 / 26.2%) or L+H* (11 / 26.2%) as their nuclear pitch accent. The large variation found in contact questions as well as in neutral understanding questions seems to reflect the well known contour variation in yes-no questions found in spontaneous speech [12].

3.2. Final boundary tones

Across all subtypes of confirmation requests analysed, a non-falling boundary tone (LH%, H%, !H%) is predominant (see Figure 3). L% mainly occurs in confirmation requests coded as (i) neutral understanding questions and as (ii) neutral perception questions.

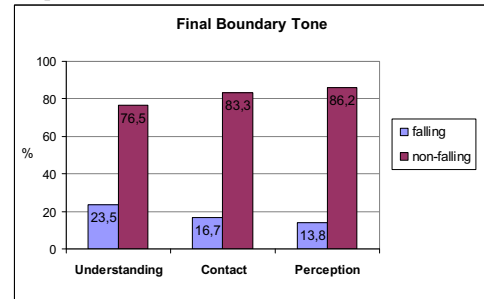


Figure 3: Percentage of falling/non-falling boundaries across subtypes of confirmation-seeking questions.

3.3. F0 height

An analysis of F0 height of non-falling boundaries shows that boundary F0 varies according to function: understanding related questions are lower than perception ones and these are lower than contact confirmation-seeking questions (either sentences or fragments). The difference between understanding and perception is significant for max(imum)_B(oundary)T(one) (Mann-Whitney Test, $U= 4569, p= .032$) and also for min(imum)_B(oundary)T(one) ($U= 2807.5, p<.0005$). See figure 4.

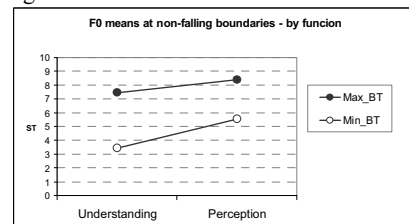


Figure 4: F0 means at non-falling boundaries across subtypes of confirmation-seeking questions

F0 range of final non-falling boundary also varies according to function: perception < understanding < contact (for perception vs. understanding $U= 3862.5, p= .006$).

With regard to F0 height on nuclear pitch accents in non-falling contours, the F0 of both H and L targets varies according to function: contact related questions are lower than understanding ones and these are lower than perception confirmation-seeking questions. The difference between understanding and perception is significant for Min(imum)_P(itc)h A(ccent) ($U= 3800, p<.0005$), although not for Max(imum)_P(itc)h A(ccent). See Figure 5.

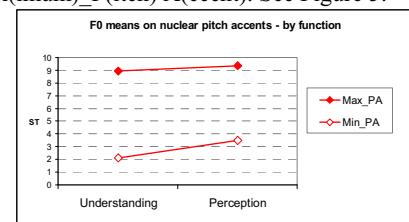


Figure 5: F0 means on nuclear pitch accents across subtypes of confirmation-seeking questions.

These data suggest that perception-seeking confirmation yes-no questions are realized in a higher register than understanding-seeking questions. This is consistent with the results obtained for English by [7].

4. Discussion and conclusion

To sum up, in our corpus, confirmation-seeking questions are not associated to a single intonation pattern; on the contrary, different intonation patterns seem to correlate with different pragmatic values. We should highlight two major results: (i) the difference between the intonation patterns of ‘understanding’ and ‘perception’ confirmation-seeking questions; (ii) the difference between the intonation patterns of ‘understanding’ confirmation-seeking questions associated with neutral and non-neutral acceptance.

‘Understanding’ and ‘perception’ confirmation-seeking questions correlate with different pitch accent types and F0 height. On the one hand, H* and L+H* mainly occur in questions seeking confirmation of perception; L*+H mainly occurs in questions seeking confirmation of understanding, even though there is no univocal correspondence between pitch accent and pragmatic value. On the other hand, perception-seeking yes-no questions seem to be uttered in a higher register than understanding-seeking questions.

The intonation difference between understanding and perception questions may be understood if we use the concept of topic, defined in terms of *aboutness* (what the sentence or the discourse is about) [13] – we particularly need the notion of D(iscourse)-topic. We may understand the intonation cues in perception oriented confirmation requests as forcing the hearer to assume a change in the topic (the D-Topic). This changes what is asked. For instance, in the case of the confirmation-seeking question in example (1), ‘understanding’ intonation cues allow to maintain “the little duck” as both the S(sentence)-topic and D(iscourse)-topic, whereas ‘perception’ intonation cues force the hearer to assume a different D-topic. In this last case, what was said becomes the D-topic and the question is recognized as an effort to confirm what the speaker actually uttered.

Another relevant result concerns intonation cues disambiguating level of acceptance in confirmation-seeking questions coded as ‘understanding’: L*+H and ^H* are dominant in the questions expressing different types of negative ‘acceptance’. Crucially, L*+H has been associated with narrow focus in interrogative questions [11] and ^H* was associated with specification or correction of given information by [2]. We suggest that confirmation requests expressing negative acceptance share some properties with contrastive / identificational focus (see [14]). The common property is *contrast*: contexts of correction are contrastive contexts; the narrow focus contexts may have a contrastive interpretation (contrastive focus). As for non-neutral (negative acceptance) understanding contexts, they can be understood as contrastive contexts: when something new is said, if accepted, it is added to the Common Ground; if the hearer does not accept it, a non-neutral confirmation-seeking question may be asked, which forces the hearer to consider the alternatives.

So it seems that these particular accents encode Contrast, which is known to be independent of focus: it may occur both associated with focus as well as with background or topic (see [15], [16] for cases of thematic contrast; see [17], [18] for contrastive topics). The facts presented in this paper are evidence in favor of this independence and further suggest that the same pitch accent which is used to encode contrast in a

constituent of a sentence may also be used to ensure that all the material in the sentence is contrasted.

This work based on spontaneous speech (child-directed), points to particular intonation cues that can disambiguate confirmation-seeking questions. The isolation of these particular cues is of considerable relevance for automatic speech processing. The results reported in this paper and those that may be obtained in the near future will also be particularly relevant to the study of language acquisition. At the present moment, we carry on a study of children’s answers trying to determine whether children’s answers are affected by the form and function of confirmation-seeking questions.

5. Acknowledgements

We would like to thank Julia Hirschberg for her helpful comments and Luísa Loura for her statistical advice.

6. References

- [1] Santos, A. L., *Minimal Answers. Ellipsis, syntax and discourse in the acquisition of European Portuguese*, John Benjamins, Amsterdam / Philadelphia, 2009.
- [2] Viana, C., S. Frota, I. Falé, F. Fernandes, I. Mascarenhas, A. I. Mata, H. Moniz and M. Vigário, “Towards a P_ToBI PAPI2007”, Workshop on the Transcription of Intonation in Ibero-Romance, Universidade do Minho, 2007.
- [3] Grice, M. and M. Savino, “Can pitch accent type convey information status in yes-no questions?”, in K. Alter, H. Pirker and W. Finkler [Eds.], *Proceedings of the ACL97 Workshop on Concept-to-Speech Generation Systems*, Universidad Nacional de Educación a Distancia, Madrid, 1997.
- [4] Grice, M., R. Benzmüller, M. Savino and B. Andreeva, “The intonation of queries and checks across languages: data from Map Task dialogues”, *ICPhS 95*, 648–651, 1995.
- [5] Rodríguez, K. J. and D. Schlangen, “Form, intonation and function of clarification requests in German task-oriented spoken dialogues”, *SemDial 2004*, Barcelona, 2004.
- [6] Edlund, J., D. House and G. Skantze, “The effects of prosodic features on the interpretation of clarification ellipses”, *Eurospeech 2005*, Lisbon, 2005.
- [7] Venditti, J. J., J. Hirschberg and J. Liscombe, “Intonational cues to student questions in tutoring dialogs”, *Interspeech 2006 - ICSLP*, Pittsburgh, Pennsylvania: 549-552, 2006.
- [8] Liscombe, J., J. J. Venditti & J. Hirschberg (submitted) “Classifying the form and function of student questions in spoken tutorial dialogs”.
- [9] Prévot, L., *Structures sémantiques et pragmatiques pour la modélisation de la cohérence dans des dialogues finalisés*. Doctoral Diss., Université Paul Sabatier, Toulouse III, 2004.
- [10] Clark, H. H., *Using Language*, CUP, Cambridge, 1996.
- [11] Frota, S., “Nuclear falls and rises in European Portuguese: A phonological analysis of declarative and question intonation”, *Probus 14*, 113-146, 2002
- [12] Mata, A. L., *Questões de Entoação e Interrogação em Português. Isso é uma Pergunta?*, MA dissertation, FLUL, 1992.
- [13] Reinhart, T., *Pragmatics and Linguistics: An Analysis of Sentence Topics*, Bloomington, Indiana University Linguistics Club, 1982.
- [14] Kiss, K., “Identificational focus versus information focus, Language”, *Journal of the Linguistic Society of America*, 74(2): 245-273, 1998.
- [15] Steedman, M., “Information Structure and the Syntax-Phonology Interface”, *Linguistic Inquiry*, 31.4: 649-689, 2000.
- [16] Braun, B. “Answers to the perception of thematic contrast and questions regarding the perception of thematic ‘non-contrast’” *Speech Prosody 2004*, 2004.
- [17] Büring, D., “Topic”, in Peter Bosch and Rob van der Sandt [Eds.] *Focus. Linguistic, Cognitive, and Computational Perspectives*, CUP, 142-165, 1999.
- [18] Büring, D., “Semantics, Intonation and Information Structure”, in G. Ramchand and C. Reiss [Eds.] *The Oxford Handbook of Linguistic Interfaces*, OUP, Oxford, 2007.