1. Introduction

In this paper, we present exploratory research on intonation in early spontaneous dialogues in European Portuguese (EP), focusing on intonation patterns in children’s affirmative answers. We explore possible correlations between intonation in children’s answers and the intonation pattern and the pragmatic value of adult’s questions, which will allow us to understand whether the choice of an intonation pattern in the answer is determined by an early sensitivity to the pragmatic value of the question, which is partly encoded by the intonation of the question. This will add to our knowledge of early intonation development.

Previous research on intonation in spontaneous dialogues has already looked at the relation between discourse function and intonation in adult’s short answers. In particular, it has been shown that intonation can distinguish answers according to function (acknowledgments/agreements, backchannels) – see Curl & Bell (2001) for yeah “agreements” and yeah “backchannels”; Benus et al. (2007) for “backchannels” and agreement short answers.

On the other hand, previous research has also shown the existence of “pitch concord” mechanisms in dialogue. These mechanisms result in “relative pitch height” matching in dialogue and are known to characterize a cooperative behavior in adult-adult dialogues (see Brazil, 1985; Couper-Kuhlen & Selting, 1996; Wichmann, 2000; Wennerstrom, 2001; Heldner, Edlund & Hirschberg, 2010 show that pitch in backchannels adapts to the pitch in the interlocutor’s utterance).

This means that the adult intonation of short answers may be expected to be affected by at least two factors: (i) the pragmatic function of the answer, which is related to the pragmatic function of the question and (ii) the pitch height in the preceding utterance. In our research, we investigate to what extent children’s short answers are also affected by these two factors.

Previous literature on early intonational development has shown that early stages of language development are characterized by a high proportion of falling contours (see Snow, 2006 and references therein and Chen & Fikkert, 2007.)

* Ana Isabel Mata (aim@fl.ul.pt), Ana Lúcia Santos (als@fl.ul.pt), FLUL/CLUL, University of Lisbon, Portugal. This work was partially funded by FCT project PTDC/LIN/66202/2006.
a.o.). On the other hand, recent work has suggested that the adult inventory of nuclear accents and boundary tones and their F0 alignment patterns are mastered very early, in certain cases before or around 2;0 – this has been suggested by Frota & Vigário (2008) for Portuguese, Vanrell et al. (2010) for Spanish and Catalan, a.o., although Chen & Fikkert (2007) do not confirm for Dutch a very early acquisition of the adult tonal inventory.

As far as F0 scaling is concerned, Frota & Vigário (2008) suggest that it is still under development at 2;0; Snow (2006) shows that accent F0 range develops at about 18 months. Given these two types of results, it is unclear whether we should expect pitch modification effects or effects of pragmatic function in the intonation of children’s short answers. However, if these effects are found, we expect age differences, namely a difference between <2;0 and >2;0.

In the next section, we report previous results on the intonation of confirmation-seeking yes-no questions in child-directed speech. In section 3 we describe the data used for the present study and in section 4 we report results on the intonation of children’s answers to those confirmation-seeking questions, pointing out that after 2;0 answers start to show convergence with the intonation pattern of declaratives in EP and correlation with subtypes of questions’ function. Finally, in section 5 conclusions are presented.

2. Previous work on the intonation of confirmation-seeking questions

In a previous study (Mata & Santos, 2010), we analyzed 301 confirmation-seeking yes-no questions uttered by three adults and answered by two children during 28h of spontaneous speech (a subset of the corpus of Santos, 2006/2009). These 301 questions correspond to all the confirmation-seeking yes-no questions uttered by the adults, that were given an affirmative answer by children and that were suitable for an acoustic analysis.

The 301 questions were coded according to three levels of action that correspond to the problem justifying the confirmation request (see Clark 1996, and also Rodríguez & Schlangen 2004, Venditti, Hirschberg & Liscombe 2006): contact, perception and understanding. 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.

The analysis of the different sets of confirmation-seeking questions shows that pragmatic/ discourse subtypes of questions affect nuclear pitch accent distribution and F0 height:

---

1 The definition of confirmation-seeking questions was assumed to be according to Prévot (2004).
(i) $H^*$ and $L+H^*$ mainly occur in perception related confirmation-seeking questions;
(ii) $L^*+H$ mainly occurs in understanding related confirmation-seeking questions;
(iii) In general, perception-seeking yes-no questions are uttered in a higher register than understanding-seeking questions (higher F0 for both nuclear pitch accent and final boundary tone).

Confirmation-seeking questions related to ‘understanding’ were also coded according to level of acceptance (neutral / non-neutral), another level of action (Clark, 1996). Non-neutral acceptance includes the suggestion of incorrectness, disbelief or surprise. Results have shown that questions expressing different levels of acceptance correlate with different pitch accents: $H+L^*$ (the most commonly used nuclear accent in information-seeking questions in EP) mainly occurs in neutral questions; in non-neutral confirmation-seeking questions we mainly find $L^*+H$ and $^H*$. It is important to notice that these pitch accents occurring in understanding questions coded as non-neutral may both be associated with the notion of Contrast (see Mata & Santos, 2010 for discussion): $L^*+H$ has been associated with narrow / contrastive focus in questions by Frota (1998/2000, 2002); $^H*$ was associated with specification or correction of given information by Viana et al. (2007).

In the present paper, we examine the intonation patterns of children’s affirmative answers to these different subtypes of confirmation requests uttered by adults (understanding, including neutral vs. non-neutral acceptance, perception, contact).

3. Data

For the present study, we use the same corpus of child-directed speech that was examined for the study of confirmation-seeking questions: 232 question-answer pairs were analyzed (<2;0 – 89; ≥ 2;0 – 143), corresponding to 77.1% of the total amount of 301 questions analyzed in Mata & Santos (2010). 69 answers were excluded (31 to understanding-seeking yes-no questions, 26 to perception-seeking and 12 to contact-seeking ones) from the 301 question-answer pairs that we had selected initially, due to background noise or speech overlapping.

Table 1. Number of answers per child < 2;0 and ≥ 2;0

<table>
<thead>
<tr>
<th>Child</th>
<th>Age</th>
<th>MLUw</th>
<th>Answers &lt; 2;0</th>
<th>Answers ≥ 2;0</th>
</tr>
</thead>
<tbody>
<tr>
<td>INM</td>
<td>1;7.6 - 2;6.19</td>
<td>1.285 - 2.265</td>
<td>68</td>
<td>80</td>
</tr>
<tr>
<td>TOM</td>
<td>1;7.14 - 2;8.9</td>
<td>1.272 - 2.979</td>
<td>21</td>
<td>63</td>
</tr>
</tbody>
</table>

All answers are grammatical affirmative answers and correspond to one of the following patterns:
- *sim* ‘yes’;
- ‘is’ / foi ‘was’ – a frozen form of the verb ser ‘to be’;
- main verb or auxiliary in a VP ellipsis structure (Santos, 2006/2009).

94% of these (one-word) short answers correspond to one syllable; the remaining 6% present two syllables.

**Intonation coding and F0 measurements.** The 232 question-answer pairs were analyzed using the WaveSurfer 1.8.5 speech analysis package ([http://www.speech.kth.se/wavesurfer/](http://www.speech.kth.se/wavesurfer/)). The intonation coding of the corpus was done manually, using the system Towards a P_ToBI by Viana et al. 2007 (see [http://www.ling.ohio-state.edu/~tobi/]([http://www.ling.ohio-state.edu/~tobi/]()). 7 out of the 8 pitch accents (H+L*, H*+L, L*+H, L+H*, H*, L*, ^H*) and all the final boundary tones (L%, H%, !H%, LH%, HL%) that are covered in that proposal were used in this task (see the schematic F0 contours in figure 1).

![Schematic F0 contours for the pitch accents and boundary tones used in the transcription of question-answer pairs, following Towards a P_ToBI. (Lines — indicate the stressed syllable.)](http://www.speech.kth.se/wavesurfer/)

For each question and answer, F0 values corresponding to high and low targets within nuclear accents and final boundary tones were measured and converted to semitones (ST). In the case of pitch accents, the maximum value of F0 (H) and the minimum value of F0 (L) were measured inside (or in the immediate vicinity) of the stressed syllable of the accented word. In the case of boundary tones, identical F0 data (maxima and minima) was extracted at the rightmost edge of questions and answers. F0 range (the difference between F0 maximum and F0 minimum values) was also extracted for both pitch accents and final boundary tones.

4. Results
4.1. Intonation of children’s answers to confirmation-seeking questions

The analysis of 232 children’s answers to confirmation-seeking yes-no questions shows that from 2;0 onwards there is a general increase of low /
falling accents ($\chi^2 (1)= 11.946, p=.001$). Whereas before 2;0 there is approximately 50% cases of high / rising - (L+)H*, L*+H - and low / falling pitch accents - (H+)L* -, after 2;0 about 70% of the answers present low / falling pitch accents (see table 2).

Table 2. Distribution of pitch accent types across affirmative answers per subtype of confirmation request and age group

<table>
<thead>
<tr>
<th>Subtypes of confirmation requests uttered by adults</th>
<th>Age Group</th>
<th>Children’s affirmative answers Pitch Accents</th>
<th>N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>High / Rising (L+)H*, L*+H</td>
<td>Low / Falling (H+)L*</td>
</tr>
<tr>
<td>Understanding</td>
<td>&lt; 2 years</td>
<td>22</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>\geq 2 years</td>
<td>21</td>
<td>38</td>
</tr>
<tr>
<td>Perception</td>
<td>&lt; 2 years</td>
<td>22</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>\geq 2 years</td>
<td>18</td>
<td>44</td>
</tr>
<tr>
<td>Contact</td>
<td>&lt; 2 years</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>\geq 2 years</td>
<td>4</td>
<td>18</td>
</tr>
<tr>
<td>N (%)</td>
<td>&lt; 2 years</td>
<td>47 (52.8)</td>
<td>42 (47.2)</td>
</tr>
<tr>
<td></td>
<td>\geq 2 years</td>
<td>43 (30.1)</td>
<td>100 (69.9)</td>
</tr>
</tbody>
</table>

We also find in answers uttered after 2;0 a general increase of falling final boundary tones (L%) – this increase is particularly visible in answers to perception and contact questions (see figure 2). The general increase of low / falling pitch accents, together with the global increase of falling boundary tones results in a clear association of these answers with a falling nuclear contour from 2;0 onwards.

![Figure 2: Percentage of falling/non-falling boundaries across affirmative answers per subtype of confirmation request and age group](image-url)
Many studies (see Snow, 2006 and references therein; Chen & Fikkert, 2007, a.o.) have suggested that there is a general tendency, in first stages of acquisition, to produce falling contours. However, the age effect that we have found in our data excludes this type of explanation. We rather interpret the increase of falling contours in affirmative answers as the result of a convergence with the intonation expected in declaratives: H+L* L% is the canonical contour of declaratives in EP.

Nevertheless, the results in table 2 also call for an analysis of answers by subtypes of questions’ pragmatic values (function). Looking at particular subcategories, we can see that answers to perception related confirmation requests are more likely to present low / falling pitch accents after 2:0 years-old ($\chi^2 (1) = 10.563$, $p=.001$). Answers to contact questions always define themselves predominantly as low / falling in both age groups, although the proportion of low/falling raises after 2:0. Finally, in the case of answers to understanding related confirmation requests, although there is an increase of low / falling pitch accents after age 2:0, the difference between the two age groups is not significant. This requires further inquiry into understanding related questions.

4.2. Intonation cues in the answers to (non-)neutral questions

Pitch accents. To understand the different behavior in answers to understanding related confirmation questions, we must recall that these questions are not all equivalent, instead we find in this group two pragmatic subtypes: neutral and non-neutral acceptance. According to our previous study (see section 2), non-neutral understanding confirmation requests associate with particular pitch accents: L*+H, ^H*; neutral ones frequently take H+L*, the most commonly used pitch accent in EP yes-no questions. When we look at answers to understanding confirmation questions taking into account the acceptance level encoded in the question, results start speaking by themselves. In fact, whereas before 2:0 we do not find a clear tendency favoring high / rising or low / falling accents in answers to either neutral or non-neutral questions, after 2:0 only answers to neutral questions converge with the general tendency to present low / falling accents (see table 3). This means that, after 2:0, it is more likely for an answer to a non-neutral question than for an answer to a neutral one to associate with a high / rising accent. The distribution differences of high / rising and low / falling accents across answers to neutral and non-neutral questions are statistically significant after 2:0 ($\chi^2 (1) = 4.664$, $p<.05$); for ages < 2:0 no correlation was found.

F0 height. An analysis of F0 height differences between answers to neutral and non-neutral understanding questions shows that F0 maxima of accents is affected: Max(imum)_P(itch) A(ccent) is higher for answers to non-neutral than for answers to neutral questions from 2:0 onwards (Mann-Whitney Test, U=
272.50, p<.05), although not before 2:0\(^2\). Moreover, as figure 3 shows, when answers are broken down by contour type, falling contours associated with non-neutral acceptance are higher than falling contours associated with neutral acceptance.

Table 3. Distribution of pitch accent types across affirmative answers to understanding related confirmation requests by acceptance level and age group

<table>
<thead>
<tr>
<th>Subtypes of confirmation requests uttered by adults</th>
<th>Age Group</th>
<th>Children’s affirmative answers Pitch Accents</th>
<th>N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>High / Rising</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(L+)H*, L*+H</td>
<td></td>
</tr>
<tr>
<td>Understanding neutral acceptance</td>
<td>&lt; 2 years</td>
<td>13</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>≥ 2 years</td>
<td>6</td>
<td>22</td>
</tr>
<tr>
<td>Understanding non-neutral acceptance</td>
<td>&lt; 2 years</td>
<td>9</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>≥ 2 years</td>
<td>15</td>
<td>(44.8)</td>
</tr>
<tr>
<td>N (%)</td>
<td>&lt; 2 years</td>
<td>22 (47.8)</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>≥ 2 years</td>
<td>21 (35.6)</td>
<td>(100)</td>
</tr>
</tbody>
</table>

These results suggest that after age 2:0 children are using two answer strategies: answers to a non-neutral understanding-seeking question may be signaled either by a non-falling nuclear contour or by a falling nuclear contour which starts higher than one associated with an answer to a neutral question.

\(^2\) An analysis of Maximum Contour produced similar results ≥2:0 (Mann-Whitney Test, U= 271.50, p<.05). As for F0 range of PA (F0max-F0min PA) in answers ≥2:0, answers to neutral present lower range than answers to non-neutral, although the difference is not statistically significant.
Examples below illustrate these answer strategies\(^3\). Intonation cues signaling non-neutral questions were associated with Contrast by Mata & Santos (2010). Apparently, around 2;0 the child starts to show some awareness of prosodic encoding of Contrast.

\[\text{Answers to non-neutral questions} \quad \text{Answers to neutral questions}\]

\text{Falling contours}

\begin{itemize}
  \item a. Child: INM
  \item b. Child: INM
  \item c. Child: TOM
  \item d. Child: TOM
\end{itemize}

\[\text{Figure 4: Falling contours in answers to (non-)neutral understanding-seeking questions by two children $\geq 2;0$.}\]

\(^3\) The figures were done using Praat and Pauline Welby’s scripts. (These scripts may draw the F0 contour in unvoiced portions of the signal.)
On the other hand, we wanted to know whether children’s answers show pitch concord effects with respect to the question. In fact, there is a correlation between F0 maxima of nuclear contours in questions and answers after age 2;0, in the case of answers to understanding related questions. For answers with a falling contour, the correlation is significant between Max(imum) P(itch) A(ccent) (as well as Maximum_Contour) of the understanding confirmation question and Max(imum) P(itch) A(ccent) of the answer ($r_s = .480$, $p$ (one-tailed) < .0005). As for answers with a non-falling contour, the correlation is significant between Max(imum) P(itch) A(ccent) of the understanding confirmation question and Max(imum) B(oundary) T(one) of the answer ($r_s = .467$, $p$ (one-tailed) < .05). Before age 2;0 no correlations were found.

These facts show that in both answer strategies mentioned above, the beginning of a falling contour and the end of a non-falling one correlate with
nuclear F0 height of questions, suggesting that the child may use mechanisms of pitch modification across turns.

Finally, we would like to show that the asymmetries in the distribution of falling / non-falling contours according to age are not due to variability in the input. We compared the number of occurrences of different pitch accent types in each subtype of confirmation request addressed by adults to children (understanding, perception, contact, neutral vs. non-neutral acceptance) before and after 2;0 years-old: the analysis reveals no significant differences. This means that, with regard to nuclear accents, a question addressed to a child before age 2;0 is similar to a question directed to a child from age 2;0 onwards. Moreover, pitch accents in children’s answers are not a simple copy of nuclear accents in the question - no correlations were found between accent types in the question and accent types in the answer. Considering understanding questions expressing neutral and non-neutral acceptance, different nuclear pitch accents associated with the different pragmatic values are produced either before 2;0 or after 2;0 (Contrast between accent types High/Rising and Low/Falling: <2;0 $\chi^2 (1)= 5.858, p<.05$ ; ≥ 2;0 $\chi^2 (1)= 14.930, p<.0001$. Contrast between nuclear accents L*+H, ^H*, H*, L+H* H*+L, L*, H+L*: <2;0 Fisher Exact Test $p<.05$ ; ≥ 2;0 Fisher Exact Test $p<.0001$.)

With regard to F0 height, an important effect of acceptance level is also observed for understanding related confirmation questions (with non-falling nuclear contours$^4$): H targets of pitch accents have higher F0 in non-neutral questions than in neutral ones; H targets of final boundaries have lower F0 in non-neutral questions than in neutral ones. A comparison based on the age group of the children to whom these questions were addressed shows that Max(imum)_B(oundary) T(one) differences between neutral and non-neutral questions are significant for both age groups (see figure 6b). However, as for Max(imum)_P(itch) A(ccent), the difference between neutral and non-neutral questions is significant after 2;0 (Mann-Whitney Test, $U= 342.50$, $p<.05$), although not before 2;0 (see figure 6a).

\[\text{Figure 6: Nuclear F0 height (F0 maxima) across understanding-seeking questions by acceptance level and age group}\]

$^4$ A non-falling boundary tone (LH%, H%, ^H%) occurs in more than 75% of the understanding questions that were analyzed and is predominant across all subtypes of confirmation requests in our data.
It is known that child-directed speech may differ from speech directed to adults in two main ways: by an exaggerated marking of contrasts (e.g. prosodic contrasts) that exist in adults speech (see Fernald & Simon, 1984, for a recent review, Payne et al., 2010, and also Kempe, Schaeffler & Thoresen, 2010) or by avoiding the use of certain structures (e.g. syntactic structures) in the early stages of the linguistic development of children (see Evers-Vermeul, 2005 for Dutch and Costa et al., 2008 for Portuguese, based on this same corpus).

The above results for adult-child question-answer pairs pointed out that the phonetic contrasts between neutral and non-neutral understanding questions may be more strongly marked for questions addressed to older children, suggesting that the use of mechanisms disambiguating the interpretation of these subtypes of confirmation questions may increase with age.

5. Conclusions

Our aim in this paper was to examine whether the intonation patterns of children’s verbal responses varies according to the discourse context and signals early sensitivity to the pragmatic interpretation of the input.

Results described in the present work, together with those discussed in a previous one (Mata & Santos, 2010), have shown that intonation cues (nuclear accent types and F0 height of both nuclear accents and boundary tones) can disambiguate the interpretation of confirmation-seeking yes-no questions in child-directed speech in European Portuguese.

In this paper, we analyzed question-answer correlations, concentrating on the intonation patterns of children’s one-word affirmative answers to confirmation-seeking yes-no questions. The analysis of affirmative answers indicates a convergence with the intonation pattern of declaratives in European Portuguese from 2;0 onwards, although not before. The answers under analysis before 2;0 suggest the existence of a fuzzy period (approximately 50% have high/rising pitch accents) followed by convergence with the adult grammar.

Moreover, we have shown that after 2;0 intonation cues in children’s answers vary according to the discourse / pragmatic context and correlate with subtypes of questions’ function: low/falling pitch accents - (H+)L* - and high/rising pitch accents - (L+)H*, L*+H -, as well as F0 height (F0 maxima within pitch accents) of answers correlate with neutral/non-neutral acceptance in “understanding” confirmation-questions.

Finally, after 2;0 there is evidence for a correlation between nuclear F0 height in questions and answers, suggesting that the child may use mechanisms of pitch modification across turns.

5 A high variability was already reported for Portuguese-speaking children’s productions before 2 years-old with regard to other prosodic parameters, namely word stress (see Correia, 2009).
This work may be seen as a first attempt to discover more about meaning and form relations in the context of adult-child spontaneous dialogues, aiming at a better understanding of intonation-discourse interactions in EP.

References


Evers-Vermeul, Jacqueline. 2005. The Development of Dutch Connectives: Change and Acquisition as Windows on Form-Function Relations. Utrecht: LOT.


