

Is prosodic development correlated with grammatical and lexical development? Evidence from emerging intonation in Catalan and Spanish*

PILAR PRIETO

ICREA-Universitat Pompeu Fabra, Spain

ANA ESTRELLA

Catholic University of Quito, Ecuador

JILL THORSON

Brown University, USA

AND

MARIA DEL MAR VANRELL

Universitat Pompeu, Fabra, Spain

(Received 11 August 2009 – Revised 16 March 2010 – Accepted 29 December 2010)

[*] The work reported in this article was presented at the International Congress for the Study of Child Language (IASCL), Edinburgh, 1–4 August, 2008, and at the XVIth International Congress of Phonetic Sciences (IcPhS), Saarbrücken, 6–10 August 2007. The authors would like to thank the audience of these conferences for their helpful comments and discussion of some of the topics dealt with in this article, and especially LI. Astruc, A. Chen, L. D’Odorico, P. Fikkert, S. Frota, C. Lleó and K. Demuth for very helpful comments. We are grateful to the action editor and two anonymous reviewers for their valuable comments on an earlier version, which have led to a significant improvement of the article. We are particularly indebted to M. Serra, S. López-Ornat and A. Ojea and M. Llinàs for generously sharing their Catalan and Spanish databases in CHILDES and granting us access to the original videotapes. We would also like to thank Y. Rose and B. MacWhinney for their help during the early stages of transcription with the *Phon* program and for developing an automatic transcription tool for Catalan and Spanish within *Phon*. We are also grateful to our colleagues A. Bonafonte and A. Moreno at the Universitat Politècnica de Catalunya for granting us access to a huge electronic dictionary containing phonetic transcriptions for Catalan and Spanish, which was the basis for the automatic transcription tool. Finally, thanks to Yoonsook Mo and Tae-Jin Yoon for help and advice on statistical measures to rate intertranscriber reliability. This research was supported by grants FFI2009-07648/FILO and CONSOLIDER-INGENIO 2010 ‘Bilingüismo y Neurociencia Cognitiva CSD2007-00012’ awarded by the Spanish Ministry of Science and Innovation and by project 2009 SGR 701 awarded by the Generalitat de Catalunya. Address for correspondence: ICREA-Universitat Pompeu Fabra – Departament de Traducció i Ciències del Llenguatge, Edifici Roc Boronat Roc Boronat 138, Barcelona, Barcelona 08018, Spain. tel: 93 2254899; e-mail: pilar.prieto@upf.edu

ABSTRACT

This investigation focuses on the development of intonation patterns in four Catalan-speaking children and two Spanish-speaking children between 0;11 and 2;4. Pitch contours were prosodically analyzed within the Autosegmental Metrical framework in all meaningful utterances, for a total of 6558 utterances. The pragmatic meaning and communicative function were also assessed. Three main conclusions arise from the results. First, the study shows that the Autosegmental Metrical model can be successfully used to transcribe early intonation contours. Second, results reveal that children's emerging intonation is largely independent of grammatical development, and generally it develops well before the appearance of two-word combinations. As for the relationship between lexical and intonational development, the data show that the emergence of intonational grammar is related to the onset of speech and the presence of a small lexicon. Finally, we discuss the implications of these results for the biological hypothesis of intonational production.

INTRODUCTION

Recent studies on prosodic development have claimed that substantial advances in the ACQUISITION OF INTONATION co-occur with more general changes in GRAMMATICAL DEVELOPMENT (SNOW, 2000; 2006; SNOW & BALOG, 2002). As SNOW (2006: 294) points out, "the milestone event in children's acquisition of expressive syntax is the appearance of two-word combinations at about 18 months of age, which coincides exactly with the dramatic growth in intonation that was observed in this and other studies". Yet some recent findings seem to contradict this hypothesis. For example, PRIETO and VANRELL (2007) recently reported that Catalan children's emerging intonation is not synchronous with grammatical development and the start of two-word combinations. The four children analyzed in that study mastered the production of a wide variety of language-specific pitch accents and boundary tone combinations well before they produced two-word utterances, regardless of the fact that the age of the start of two-word production was 1;6 for two of the children and 2;0 for the other two. The fact that these children had an important knowledge of intonational grammar well before their first two-word utterances casts doubt on the hypothesis that children's development of grammar coincides in time with the development of intonation and suggests that the development of intonational grammar occurs before grammatical development. Similarly, FROTA and VIGÁRIO (2008) found that a European Portuguese child acquired the inventory of pitch accents and boundary tones in an adult-like way at 1;9, with the emergence of such contours as early as 1;5. For this European

Portuguese child, intonational development occurred five months before the onset of the two-word stage, which for this child was 2;2.

On the other hand, recent studies on the acquisition of Dutch and European Portuguese intonational patterns have found that intonational development is correlated with an INCREASE IN VOCABULARY SIZE (Chen & Fikkert, 2007; Frota & Vigário, 2008). In Chen and Fikkert's (2007: 315) study, this correlation was found in three children aged between 1;4 and 2;1. They showed that all children mastered the basic inventory of the boundary tones and nuclear pitch accent types at the 160-word level, and the set of non-downstepped prenuclear pitch accents at the 230-word level. In Frota and Vigário's (2008) study, the monolingual toddler acquired the adult-like inventory of pitch accents and boundary tones at 1;9, which coincided in time with a vocabulary size of more than 20 words. Similarly, Vihman and DePaolis (1998) and Vihman, DePaolis and Davis (1998) found that English and French infants began to use fundamental frequency (or *f₀*) patterns consistent with the adult language at the 25-word point. This large discrepancy in lexicon size between the Dutch and the Portuguese, French and English children at the time of the intonational boost calls for a deeper understanding and investigation of the relationship between intonational and lexical development.

The first purpose of this investigation is to describe the intonational properties of early utterances in Catalan and Spanish. Specifically, we address the following questions: (1) When do Catalan and Spanish children acquire their basic intonation patterns and the inventory of nuclear pitch accent configurations? (2) Do the children master the alignment and scaling properties of pitch accents and boundary tones in the language from the beginning? This work is one of the first investigations of early intonation patterns of Catalan- and Spanish-acquiring children and it enlarges the empirical coverage of intonational development in Romance languages (Lleó, Rakow & Kehoe, 2004; Lleó & Rakow, 2011, for Spanish; Prieto & Vanrell, 2007, for Catalan; Astruc, Prieto, Payne, Post & Vanrell, 2009, for Catalan and Spanish; Frota & Vigário, 2008, for European Portuguese; D'Odorico & Carubbi, 2003; D'Odorico & Fasolo, 2009, for Italian). The empirical basis for this investigation is an extensive longitudinal audiovisual corpus consisting of the transcribed speech of four Catalan children coming from the Serra-Solé corpus on Catalan available in CHILDES, and of two Peninsular Spanish children (the Llinàs-Ojea corpus and the López-Ornat corpus in CHILDES).

The second aim of this investigation is to assess whether the mastery of a number of intonation patterns by Catalan and Spanish children is correlated with grammatical and lexical development. We are interested in analyzing the temporal relationship between grammatical, lexical and intonational development across children and languages. Following recent work on

prosodic development, our hypothesis is that precocious expression of intonation patterns will not necessarily be correlated in time with syntactic and lexical developmental trajectories. Instead, the intonation patterns might only be an early indicator of language development such that prosody might drive lexical and grammatical development also in production. Studies on infant perception have revealed that the prosodic analysis of the speech signal may allow infants to start acquiring the lexicon and syntax of their native language and thus that prosody serves as a ‘guide’ for lexical and syntactic acquisition (Christophe, Guasti, Nespor, Dupoux & van Ooyen, 1997; Christophe, Gout, Peperkamp & Morgan, 2003; Nespor, Guasti & Christophe, 1996; among many others). The main purpose of this article is to investigate whether prosody drives the development of syntax and lexical development, and thus that prosodic development would come before grammatical and lexical development. Yet it is still an open question whether prosodic production abilities in children are paced in some way with grammatical and lexical development, even if they are discontinuous in time.

Intonation in early child speech has traditionally been analyzed from a holistic perspective. In general, the whole utterance (or the final part of the utterance) has been the unit of analysis and the contour has been described in terms of its overall rising or falling shape (see Snow, 2006, and Snow & Balog, 2002, for a review). Even though this approach has proven to be useful, some researchers have started to successfully apply the Autosegmental Metrical framework (henceforth AM framework) to investigate the early intonation patterns in child speech (see Prieto & Vanrell, 2007, for Catalan; Chen & Fikkert, 2007, for Dutch; Frota & Vigário, 2008, for European Portuguese). As is well known, the AM model (Beckman & Pierrehumbert, 1986; Gussenhoven, 2004; Jun, 2005; Ladd, 2008; Pierrehumbert, 1980; among others) has quickly become the most widely used phonological framework for analyzing intonation. In our view, the use of the AM model in early acquisition can offer a more fine-grained tool to investigate how children learn the language-specific inventory of phonologically distinct intonation contours of the target language. Given recent reports that *fo* association patterns are attained by children very early in production (see Astruc *et al.*, 2009; Kehoe, Stoel-Gammon & Buder, 1995; Prieto & Vanrell, 2007), we will assess whether an AM analysis in terms of the inventory of Catalan and Spanish adult pitch accents and boundary tones can be successfully used to transcribe early intonation contours produced by Catalan and Spanish children.

To evaluate the claim of early mastery of intonational grammar, we assessed the phonetic realization of intonation contours together with the children’s pragmatic intentions. To do this, we coded the data for sentence type and for communicative intent, basing our description on the speech act

theory (Austin, 1962; Searle, 1969) and on the application of this theory to the analysis of early utterances in children's speech (Dore, 1973; 1974; 1975; and more recently Ninio, 1992; Ninio, Snow, Pan & Rollins, 1994).

The article is organized as follows. First, we describe the Catalan and Spanish corpus materials and the methodology used for the intonational analysis of the data. Second, we present the results of the study, analyzing the development of each child along with a qualitative and quantitative analysis at both the one-word and two-word stages. Finally, we conclude with a discussion on the connection between prosody and grammatical and lexical development and we discuss the implications of the results for the analysis of prosodic development.

METHOD

Participants

The empirical basis for this study is an extensive longitudinal corpus consisting of the transcribed speech of four Catalan children (Gisel·la, Guillem, Laura and Pep) and two Spanish children (Irene and María). The Catalan data comes from the Serra-Solé corpus and the Spanish data from the Ojea corpus and López-Ornat corpus, all of which are available on the CHILDES website. The Catalan children and both of the parents of these children used Central Catalan almost exclusively in their family context (they all are from Barcelona, Spain).¹ The Spanish children and both of the parents of these children used the Northern Peninsular Spanish variety (specifically from Gijón and Madrid, Spain) in the home exclusively.

Materials

Each child was videotaped on a monthly basis approximately from the start of the use of 25 words or before that (between 0;11 and 1;8, depending on the child) up until four years of age.² Data was collected following a naturalistic design, that is, spontaneous situations were recorded at home in everyday situations with one parent and the researcher. The typical activities included reading a picture book, playing with toys, eating, etc. For

[1] Also, while none of the Catalan children are bilingual with Spanish, they do have slightly varying degrees of contact with the Spanish language outside of the home environment due to exposure from television, daycare, friends of the family, neighbors and other day-to-day events.

[2] The only exception to the 25-word start is the Spanish child María. Yet even though the recordings of María start with a use of 50 words, we think that it is important that she is part of this study. First, her data allow us to analyze her intonation contours at the 50-word level and check whether her intonational inventory fits the general predictions. Second, we can check her command of the different types of contours included in her inventory as well as her intonational development over time.

TABLE 1. *Summary of the Catalan and Spanish data : ages analyzed, number of sessions and number of utterances for each of the children in the study*

	Age	# of sessions	# of meaningful utterances
Sp_Child			
Irene	0;11.1-2;4.13	18	1361
María	1;7-2;4	9	1377
Cat_Child			
Gisella	1;7.14-2;4.25	8	890
Guillem	1;1.29-2;4.24	14	1079
Laura	1;7.20-2;4.11	7	490
Pep	1;1.24-2;4.1	14	1361
Total		70	6558

Catalan, the data was transcribed in orthographic form by a team directed by Miquel Serra and Rosa Solé, and is available on the CHILDES website (MacWhinney & Snow, 1985). For Spanish, the data was also transcribed in orthographic form and is available under the Llinàs-Ojea and López-Ornat corpora in CHILDES. Table 1 presents a summary of the data used for this study.

Table 1 lists the name of each child, their age range analyzed, the number of sessions, and the total number of meaningful utterances analyzed for each child. ‘Sp_Child’ denotes the Spanish children and ‘Cat_Child’ denotes the four Catalan children. The total number of utterances analyzed was 6558. Note that the age range analyzed is different for each child. Our data analysis spanned from the beginning of the recording sessions (generally before the 25-word point) up until past the start of the two-word utterance period, which is set to 2;4 for all children.

Corpus annotation

After digitizing the original videotapes for compatibility with *Phon* (Rose *et al.*, 2006), we segmented and phonetically transcribed the recorded data for the six children using this software.³ In this first stage, all utterances spoken by the children were segmented, including speech-like utterances such as vocalizations, cries or whisperings, but only meaningful utterances were analyzed.

The target meaningful utterances were transcribed pragmatically and prosodically by the authors. In landmark reviews of developmental

[3] We would like to thank M. Serra, S. López-Ornat and A. Ojea and M. Llinàs for generously sharing their Catalan and Spanish databases and granting us access to the original videotapes.

intonation studies, Crystal (1973; 1986) argued that children's intentions need to be assessed independently from prosody (see also Snow & Balog, 2002, for a review). For this investigation, we analyzed prosodic and pragmatic information separately to try to minimize the interaction between the two types of information. While pragmatic coding (that is, the children's intentions and the characteristics of the speech act) was performed by using video files with *Phon* (thus with access to the discourse context and the audio files), prosodic coding was performed using *Praat* (Boersma & Weenink, 2009), with no access to discourse context and visual and gestural information.⁴ In the following subsections, we explain the main rationale behind the pragmatic and prosodic codings.

Pragmatic coding

In order to assess whether children have an early command of intonational grammar, it is important to assess the phonetic realization of intonation contours together with the children's pragmatic intentions. To perform the pragmatic analysis, we based our description on the speech act theory (Austin, 1962; Searle, 1969), according to which two expressions can give rise to a complex speech act exclusively when they have one, and only one, illocutionary force.

For the pragmatic coding, on a first pass we judged each utterance to be meaningful or non-meaningful. Following Snow (2006), meaningful utterances were identified on the basis of four criteria: (1) some phonetic relation to an adult-based word; (2) appropriate use in context; (3) consistency; and (4) the parent's confirmation that the child's utterance was meaningful. Imitated utterances were also transcribed, but are not reported in this article.

After this first selection was performed, each meaningful utterance was assigned two semantic labels: (1) sentence type, according to the following possibilities – exclamatives, commands, interrogatives, requests, statements, vocatives; and (2) a semantic label based on the basic speech act primitive labels established originally by Dore (1975). Table 2 shows the quantitative distribution in our data of the six sentence types used for the pragmatic labeling. The results show that statements were by far the most frequently produced type of utterance by all of the children in both languages. Yet even though the majority of the utterances recorded were statements, there were also a variety of sentence types.

Different researchers have shown that early child speech can be successfully analyzed by using a set of basic speech acts that express a set of

[4] The reader can access the online Phon databases made for the six children at: <http://prosodia.upf.edu/phon>.

TABLE 2. *Number of utterances analyzed by sentence type for the six children*

Sentence-Types	CATALAN				SPANISH		Type Totals
	Gisel-la	Guillem	Laura	Pep	Irene	María	
Exclamatives	86	116	69	32	45	45	393
Commands	15	66	41	46	62	62	292
Interrogatives	97	124	73	7	78	78	457
Requests	36	100	19	44	106	122	427
Statements	652	634	268	1149	975	975	4653
Vocatives	4	39	20	83	95	95	336
<i>Total</i>	<i>890</i>	<i>1079</i>	<i>490</i>	<i>1361</i>	<i>1361</i>	<i>1377</i>	<i>6558</i>

communicative functions that take into account the child's pragmatic intentions (Dore 1973; 1975; and more recently Ninio, 1992; Ninio *et al.*, 1994). In our data, we used a set of labels that were intended to cover the underlying intention of the child. Transcribers judged whether an utterance had at least one illocutionary force on the basis of their perception of the communicative context, given their assessment of the situation through the video files. The video files allowed the coders to evaluate both pragmatic and gestural information, as well as the adult's reactions. The labels we used were the following: emphasis, surprise, obviousness, insistence, confirmation, request and complaint.⁵ Those labels were chosen on the basis of previous literature and of our experience in coding pragmatic meaning, and were applied to different sentence types. For example, emphasis was applied to all sentence types; insistence to vocatives, statements, requests and complaints; obviousness to statements; confirmation to *yes/no* questions; and surprise to exclamative utterances.

Prosodic coding

As mentioned before, we conducted our intonational analysis within the AM framework (Beckman & Pierrehumbert, 1986; Jun, 2005; Ladd, 2008; Pierrehumbert, 1980; among others). In the AM framework, the *fo* contour of an utterance is described as a sequence of high (H) and low (L) tones, with an additional mid tone in certain languages. The tones are of two kinds, pitch accents and boundary tones. Pitch accents are tonal events that are associated with the metrically prominent syllables in a sentence, and they can be either monotonal (e.g. H*, L*) or bitonal (e.g. L+H*, L*+H,

[5] These labels were used only when they appeared in the data, meaning that many sentence-type codings do not have a corresponding 'intention' label. For example, in the case of information-seeking questions, no corresponding intention label was used. That is why we do not present a quantitative description of these codings.














H+L*, among others). The starred tone is usually realized on the stressed syllable. Boundary tones are tonal events that are associated with the edges of prosodic phrases. They can be high (H) or low (L). The boundary tones associated with the right edges of intonational phrases (IP) are marked with a ‘%’ sign following the tone (e.g. H%, L%). An intonational phrase can have more than one pitch accent, and the final one is usually referred to as the nuclear pitch accent; the rest of the pitch accents are referred to as the prenuclear pitch accents.

The same transcriber performed both the pragmatic and prosodic codings for the same child. Each meaningful utterance was annotated for the following fields: (1) orthographic transcription; (2) prosodic transcription in the Catalan or Spanish versions of the Tones and Break Indices model, ToBI (Cat_ToBI: Prieto, Aguilar, Mascaró, Torres-Tamarit & Vanrell, 2009; Aguila, de-la-Mota & Prieto, 2009a; Prieto, in press; Sp_ToBI: Estebas-Vilaplana & Prieto, 2010). In this study, we will mainly concentrate on the description of nuclear pitch accents plus boundary tone combinations found in the data, that is, nuclear pitch configurations. In both Catalan and Spanish, the rightmost member of a prosodic phrase receives the nuclear pitch accent, that is, the most prominent accent within the phrase. Nuclear tonal configurations are an important part of intonation contours, and are key elements in the expression of a variety of pragmatic meanings in discourse. Table 3 presents a summary of the commonly occurring nuclear pitch configurations in adult Catalan.⁶ Each tune is represented by a schematic contour in the first column, followed by the Cat_ToBI label, and a possible pragmatic context where it is found. In the schematic contours, the shaded box represents the stressed syllable. For a more comprehensive description of the intonational phonetic form and pragmatic function of each of the contours, see Prieto (in press).

Table 4 presents a summary of the commonly occurring nuclear pitch configurations in adult Spanish (for a more comprehensive description, see Estebas-Vilaplana & Prieto, 2010; Aguilar, de-la-Mota & Prieto, 2009b). As we can see by comparing Tables 3 and 4, there is a great deal of overlap in the inventory of nuclear pitch configurations in Catalan and Spanish, even though the pragmatic meanings of some of the contours are different. The main differences between the phonological inventory of nuclear pitch configurations in the two languages are related to the semantic scope of some nuclear configurations: (1) while H+L* L% is a possible intonational contour of an information-seeking *yes/no* question in Central Catalan, in














[6] The reader can access both the *Cat_ToBI* and the *Sp_ToBI Training Materials*, together with audio files and exercises, at: http://prosodia.upf.edu/cat_tobi/ (Cat_ToBI) and http://prosodia.upf.edu/sp_tobi/ (Sp_ToBI).

TABLE 3. *Schematic representation of commonly used nuclear pitch configurations in Catalan, the Cat_ToBI label, and one of the common pragmatic functions (taken from Prieto, in press)*

Schematic contour	Cat_ToBI transcription	Context
	L* L%	Broad focus statement
	H+L* L%	Information-seeking <i>yes/no</i> question
	H* L%	<i>Wh</i> -question
	L* HH%	Information-seeking <i>yes/no</i> question
	L+H* HH%	Invitation/imperative <i>yes/no</i> question surprise echo question
	L+H* L%	Narrow focus statement, exclamative Imperative
	L* HL%	Statement of the obvious Soft request Disapproval statement Insistent request
	L+H* !H%	Vocative chant
	L+H* HL%	Vocative (attention request)
	L+H* LH%	Counter-expectational echo question
	L+H* LHL%	Insistent request
	L+H* LM%	Statement of the obvious Soft request
	L+H* !H%	Uncertainty statement

Spanish it is not used as an information-seeking question, but rather as a seldom-used confirmation-seeking question; (2) while L+H* HH% is mainly used as an invitation/imperative *yes/no* question in Catalan (with a

TABLE 4. *Schematic representation of commonly used nuclear pitch configurations in Peninsular Spanish, the Sp_ToBI label and one of the common pragmatic functions (adapted from Estebas-Vilaplana & Prieto, 2010)*

Schematic contour	Sp_ToBI transcription	Context
	L* L%	Broad focus statement
	H+L* L%	Confirmation-seeking question
	H* L%	Wh-question
	L* HH%	Information-seeking yes/no question
	L+H* HH%	Information-seeking yes/no question, surprise echo question
	L+H* L%	Narrow focus statement, exclamative Imperative
	L* HL%	Statement of the obvious Soft request Disapproval statement Insistent request
	L+H* !H %	Vocative chant
	L+H* HL%	Vocative (attention request)
	L+H* LH%	Counter-expectational echo question Invitation yes/no question
	L+H* LHL%	Insistent request
	L+H* L!H%	Statement of the obvious Soft request
	L+H* !H%	Uncertainty statement

nance of ‘obliging disposition’), in Spanish it has a wider scope and it can even be used as an information-seeking question. Even though we noted these differences, the pragmatic coverage of these intonation contours needs

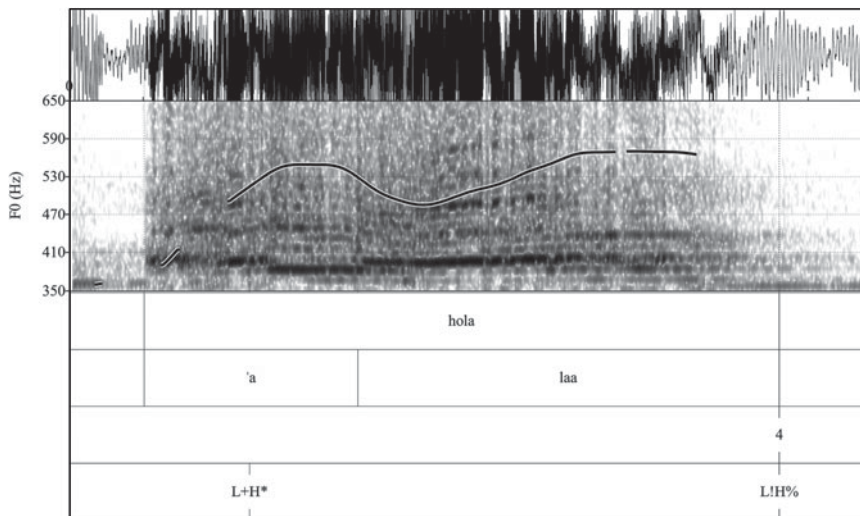


Fig. 1. Waveform display, spectrogram, fo contour and prosodic labeling of the utterance *hola* ‘hello’ produced by Guillem at 1;4:26.

to be further investigated in the two languages, but this is out of the scope of this article.⁷

Figure 1 shows a sample of the orthographic and prosodic transcription performed with the utterance *hola* ‘hello’ produced by Guillem at 1;4:26 with the meaning of a soft request. Phrase breaks are transcribed in the third horizontal tier (using phrase break number 3 and 4 to indicate the end of an intermediate phrase and the end of the intonational phrase respectively), and pitch accents and boundary tones are transcribed in the fourth, while the orthographic transcription appears in the first and the phonetic transcription on the second. In this case, the intonation produced is that of an insistent request consisting of a rise in pitch during the stressed syllable (L+H*) followed by a complex boundary tone L!H%. Finally, whenever the transcriber could note obvious differences between the adult fo contours and the children’s this was noted in a separate tier.⁸

An inter-transcriber reliability test was conducted with a subset of our data. A total of 80 utterances from the children’s databases were randomly selected by one of the authors, taking into account that all children and ages

[7] See Thorson *et al.* (2009) for a deeper investigation of the use of interrogative contours in Catalan and Spanish child speech and child-directed speech.

[8] For example, one of the phenomena that was frequently annotated in the data was the presence of fo mid tone (instead of a L% tone, marked as E% for error in our data) which typically appears at the end of statement intonation contours, and which does not appear in adult speech.

were uniformly represented. After this, the three transcribers of the corpus labeled the target utterances using the Cat_ToBI and Sp_ToBI systems. A comparison of the tonal transcription across the three transcribers reveals a 77% consistency in pitch accent and boundary tone decisions. The agreement on the choice of pitch accent is 89% and of boundary tones is 65%. In addition to the transcriber-pair-word analysis, the kappa statistic was also obtained (Randolph, 2008). This measure calculates the degree of agreement in classification over that which would be expected by chance and its value can range from -1.0 to 1.0 , with -1.0 indicating perfect disagreement below chance, 0.0 indicating agreement equal to chance and 1.0 indicating perfect agreement above chance. The main difference between the pairwise agreement measure and the kappa statistic is that the latter takes into account the number of possible categories while the former does not. Since there were three raters in our study, the Fleiss' kappa statistical measure was used (Yoon, Chavarria, Cole & Hasegawa-Johnson, 2004; Yoonsook, Cole & Lee, 2008). Other kappas such as Cohen's kappa only work when testing the agreement between two transcribers. The fixed marginal kappa statistic obtained for the choice of pitch accents and boundary tones was of 0.70 and 0.52 , respectively. While the choice of pitch accents has a kappa statistic of 0.70 , indicating that those categories were reliably labeled, the choice of boundary tones has a lower reliability measure. This is probably due to the fact that raters have to choose between many different combinations and they must face decisions about the distinction between an L% boundary tone and an undershot boundary tone (marked as E% in our data). In general, though, with a 77% agreement we can be moderately confident about the reliability of the transcriptions, as during the transcription process we met regularly to transcribe and to discuss transcription decisions.

RESULTS

Mean Length of Utterance

One of the most widely used indices of language development and grammatical complexity is the Mean Length of Utterance in morphemes (MLUm) or words (MLUw). For this study, we calculated the MLUw of each child using the 'mlu' command in CLAN. Figure 2 shows the MLUw for each of the sessions (represented on the x-axis), for each child. It is interesting to note that children display great variation regarding the time they reach an MLUw level of 1.5 , the number we will refer to when pinpointing the ESTABLISHED onset of the two-word period. Note that MLUw counts may drop a bit in-between certain sessions, possibly because the child was not as talkative and cooperative in some of the sessions. Yet for us the important thing is that the child reaches the critical MLUw level

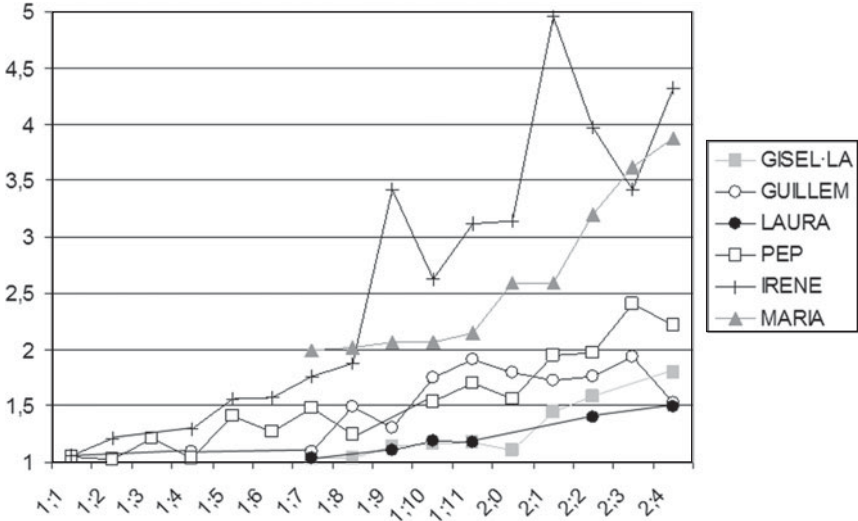


Fig. 2. Measures of Mean Length of Utterance in words for each of the sessions, for each child.

of 1;5 at a given point in time (which means that half of the utterances uttered by the child in this session were two-word utterances). In essence, we are probably underestimating when they reach these points, not overestimating. The graph shows that while Pep, Guillem and Irene all reach an MLUw level of 1;5 between the ages of 1;5 (Pep and Irene) and 1;8 (Guillem), Laura and Gisela do not reach this level until six months later or more (around 2;1). In the case of Maria, her data begins when she is 1;7 and she has already reached an MLUw of 2; this means that we will have to limit her analysis to her development after the onset of the two-word period.

The natural dual distribution of the data makes it possible to test whether there is a sound correlation between grammatical and intonational development (Snow, 2000; 2006; among others). Specifically, we will test how the MLU results for each child correlate with the acquisition of distinct nuclear configuration types (see ‘Quantitative results’ below). If Snow’s hypothesis is correct, we would expect to see a close correlation between the two measures across the six children.

Lexical development

In our data, vocabulary size was computed with the ‘freq’ command in CLAN, that is, by listing the number of unique recorded words per session.

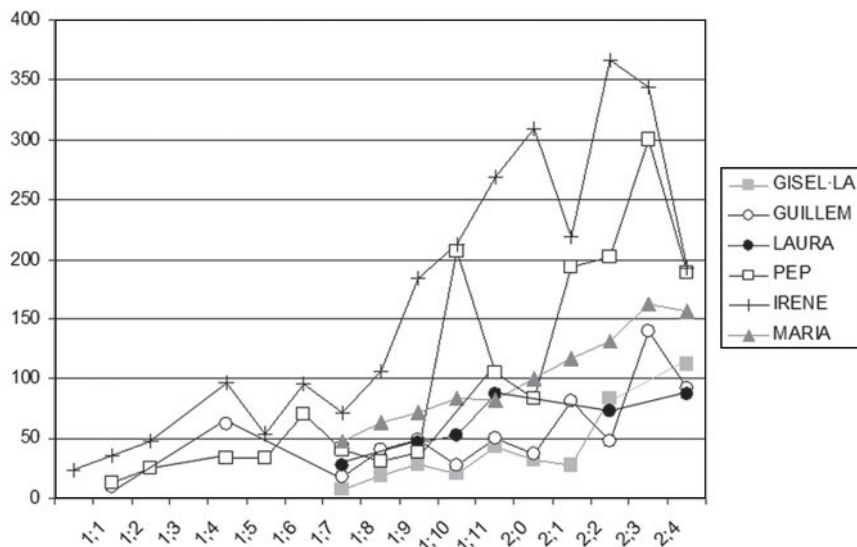


Fig. 3. Number of distinctive word types for each of the sessions, for each child.

Figure 3 shows the number of distinctive word types found for each of the sessions (shown on the x-axis), for each child. The definition of the 25-word point is the same as the one proposed by Vihman *et al.* (1998) and DePaolis, Vihman and Kunnari (2008), that is, the first month in which the child used 25 or more identifiable adult-based words spontaneously in one half-hour session. The data in Figure 3 show that, similarly to the MLU data, Pep, Guillem and Irene all reach a vocabulary size of 25 words between 0;11 (Irene) and 1;6 (Guillem). On the other hand, Laura and Gisel·la do not reach this lexicon size until they are 1;8 (Laura) and 2;0 (Gisel·la). It is important to note that even though the lexical counts fluctuate across sessions (possibly due to the child's behavior in a given session), we assume that if a child uses 25 words in a given session this is an indication that he or she has reached the 25-word point.

The data in Figure 3 show that the children's lexicon size data pattern differently from the MLU data presented in Figure 2. While Irene reaches a lexicon size of 100 words at 1;4, Pep does not reach this level until he is 1;11, and the other children not until months later, at 2;4. It is interesting to note that while Guillem gets to the two-word stage quite early (five months before Gisel·la), he patterns with them in his lexicon size, which does not get to be 100 words until he is 2;4. This seems to be a clear indication that the lexicon size and grammatical complexity measures are not strictly correlated in development.

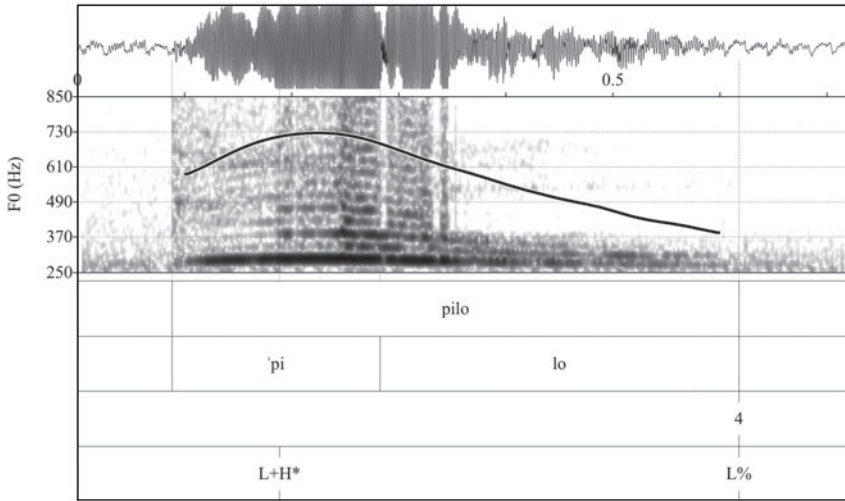


Fig. 4. Waveform display, spectrogram, fo contour and prosodic labeling of the utterance *pilota* ‘ball’ produced by Pep at 1;2:3.

Qualitative results

This section examines in a qualitative way the intonational development of all children both at the one-word and at the two-word stages. This section can be regarded as an initial overview of the data before the quantitative analysis is performed. The initial focus of the analysis will be on Guillem, Pep and Irene, the three children who produce two-word combinations stably at around 1;5 (Pep and Irene) and 1;8 (Guillem). For this part of the analysis, Maria could not be analyzed due to lack of data before the onset of the two-word period.⁹ In general, the intonational analysis reveals that all children begin to use a handful of intonational contours at the onset of the one-word period. In the case of Guillem, Pep and Irene, they produce these contours between 1;1 and 1;3.

In our data, the most widely used contour is the statement, used as a way to designate an object or as a response to a question. Among the statements, the most common nuclear pitch accent and boundary tone configuration is L + H* L%. The alignment properties of the L + H* pitch accent and L% boundary tones were largely mastered early in the intonational development of these three children. For example, Figure 4 shows the waveform, the spectrogram, and the fo contour of the utterance *pilota* ‘ball’ produced by

[9] Note that she started to be recorded when she already produced two-word combinations and eight different types of nuclear configurations (see ‘Quantitative results’ below).

Pep at 1;2·3.¹⁰ This was Pep's answer to the question by his mother *Què és això?* 'What is this?' As the fo pitch track shows, the start of the rise of the L+H* pitch accent coincides with the beginning of the stressed syllable; the end of the rise (of the fo peak) coincides with the end of the stressed syllable, and, after that, the fo falls in the post-tonic syllable.

The acquisition of word stress is very important for the development of intonation, as the intonational movements are 'anchored' in metrically strong syllables. We reported virtually no stress placement errors for any of the children. Importantly, the alignment properties of the L+H* L% nuclear configuration are largely mastered: the rise of the L+H* pitch accent starts to rise at the beginning of the syllable, and it ends towards the end of the syllable; after that, the fo falls in the post-tonic (see also Kehoe *et al.*, 1995, Astruc *et al.*, 2009, Vanrell, Prieto, Astruc, Payne & Post, 2010, for similar findings). As for the tonal scaling of tonal targets, it was noticed during the initial analyses of the data that the target L% boundary tone was not always rightly produced in all of the statements. The L% boundary tone was realized as a mid tone by the children, and not as the target low tone found in adult speech. The mid realizations of L% boundary tones were marked perceptually and an E% boundary tone was used, standing for error. Even though these contours were not used in the general quantitative analysis of the data, there is a progressive longitudinal decrease in the L% boundary tone scaling errors (the E%) as the children mature. For example, Irene begins with scaling errors in 80% of the data. Over time the general percentage decreases, with the error rate at 41% at 1;7 and disappearing almost completely to 0% by age 2;0.¹¹

In our data, there are examples that show an adult-like use of pitch accent range, which develops very fast in the use of focal accents. For example, Guillem, Irene and Pep use a wider pitch accent range to express emphasis or focus, as in the case of the emphatic or imperative utterance *Laia, Laia* 'proper name' uttered by Pep at 1;2·28 (see Figure 5), while trying to desperately catch his sister's attention. Again, alignment is target-like, with the L target aligned with the onset of the stressed syllable and the H peak aligned with the end of the stressed syllable.

[10] As noted by one of the reviewers, cross-linguistic findings in the literature suggest that children should start with a form like [ˈlota] for *pilota* 'ball' (analogous to English [ˈnana] for *banana*) (see Prieto, 2006, for an analysis of early truncation patterns in Catalan and Spanish). Instead, Pep produces [ˈpilo] instead of [ˈlota] for *pilota* 'ball'. This can be traced back to the fact that [ˈpilo] or [ˈpelo] are very common ways of truncating this word both in adult Catalan and Spanish, respectively. Thus, arguably the adult word that Pep hears is this one and the child is not really truncating the sequence.

[11] The issue of the misproduction of target fo tonal scaling at the end of statements has been investigated in a quantitative way by Vanrell *et al.* (2010).

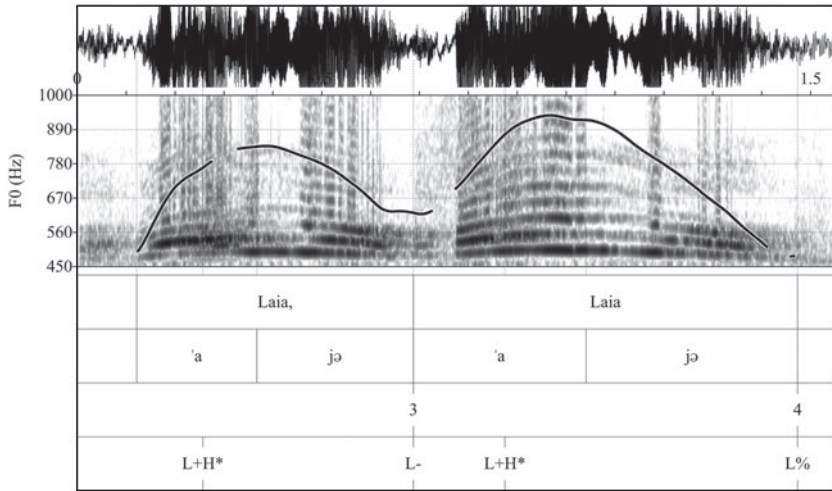


Fig. 5. Waveform display, spectrogram, fo contour and prosodic labeling of the utterance *Laia, Laia* 'proper name' produced by Pep at 1;2:28.

Another contour produced by the three children is the 'calling contour' or 'stylized call or chant', which is phonetically realized with a rising accent on the accented syllable L+H* followed by a falling-rising movement L!H% (see the utterance *hola* 'hello' produced by Guillem at 1;4:26 in Figure 1). This contour is produced with other 'chanted' utterances such as the typical pattern *ja està* 'all done'.

The precocious development of intonation during the one-word period is demonstrated by the appearance of complex boundary tones at the end of this stage. Guillem produces the complex nuclear pitch contours L+H* L!H% and L+H* HL% well before the production of two-word combinations at 1;8. For example, Figure 6 shows the intonation pattern of the utterance *papá!* 'daddy!' produced by Irene at 1;4:16. This contour is a calling contour that has the function of requesting the attention of Irene's father. It is phonetically realized with a rising pitch accent on the accented syllable (L+H*) plus a complex HL% boundary tone (cf. also Figure 1). The final boundary tone L% is not realized at the target L level but at a higher level.

At the two-word period, the three children start producing a variety of tunes to express request, discontent or insistence, patterns which are especially complex in Catalan, as well as interrogative utterances. For example, one of the disapproval contours in adult Catalan is produced with a nuclear accent L* followed by a complex HL% boundary tone. Figure 7

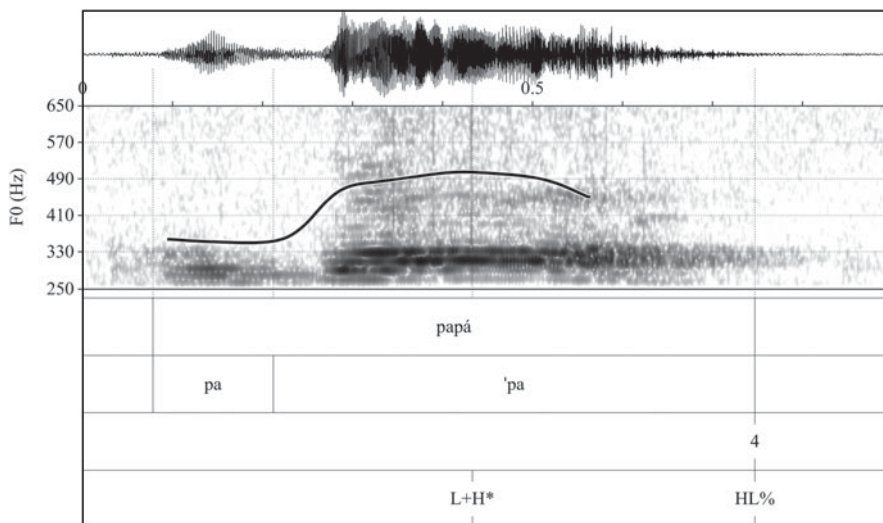


Fig. 6. Waveform display, spectrogram, fo contour and prosodic labeling of the utterance *papá!* ‘daddy!’ produced by Irene at 1;4:16.

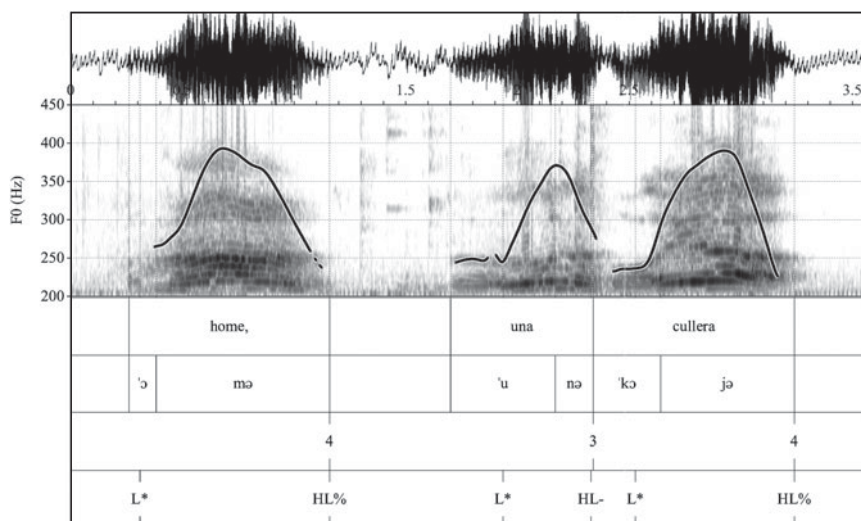


Fig. 7. Waveform display, spectrogram, fo contour and prosodic labeling of the sequence *home, una cullera!* ‘man, a spoon!’ uttered by Pep at 1;8:0.

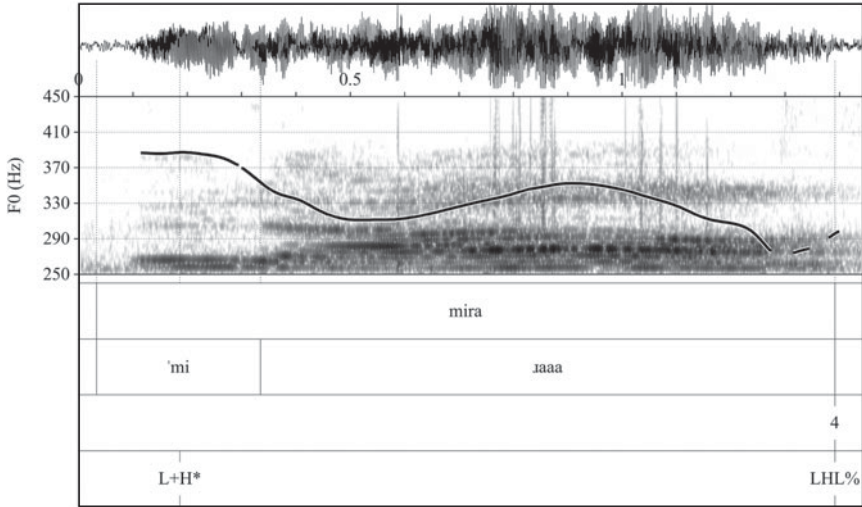


Fig. 8. Waveform display, spectrogram, fo contour and prosodic labeling of the sequence *mira* ‘please take a look’ uttered by Guillem at 1;11.13.

shows the first production of this contour by Pep: [‘ɔmə, ‘una ‘kɔʝə] *home, una cullera!* ‘man, a spoon!’

The example in Figure 7 demonstrates that the child Pep at age 1;8 is capable of successfully producing the complex tune–text association patterns that characterize some fo contours: the child associates the tone L* to the three accented syllables (*home* ‘man’, *una*, and *cullera* ‘a spoon’), and associates a complex HL% boundary tone with the post-accentual syllable.

Another example of an especially complex intonation pattern is the insistent request shown in Figure 8. Insistent requests in Catalan can be expressed through an intonation contour that consists of a L+H* pitch accent followed by a complex boundary tone sequence LHL%. The production of this contour demonstrates that relatively early Guillem has an outstanding control over the complex alignment of edge tones.

For the three children, interrogative utterances appear in the two-word period. Figures 9 and 10 show examples of Irene producing information-seeking interrogative utterances with tonal nuclear configurations of L* HH% on the phrase *otra vez?* at 1;6.16 and *puedo dar la vuelta?* at 1;11.13.

Similarly, the analysis of the intonation contours produced by Gisel·la and Laura reveal that there is a great increase in the use of intonation well before they start using two-word combinations (Gisel·la at 2;1 and Laura at 2;3; see Figure 2). By this time both produce statements and a variety of exclamative, imperative and interrogative intonation contours in an

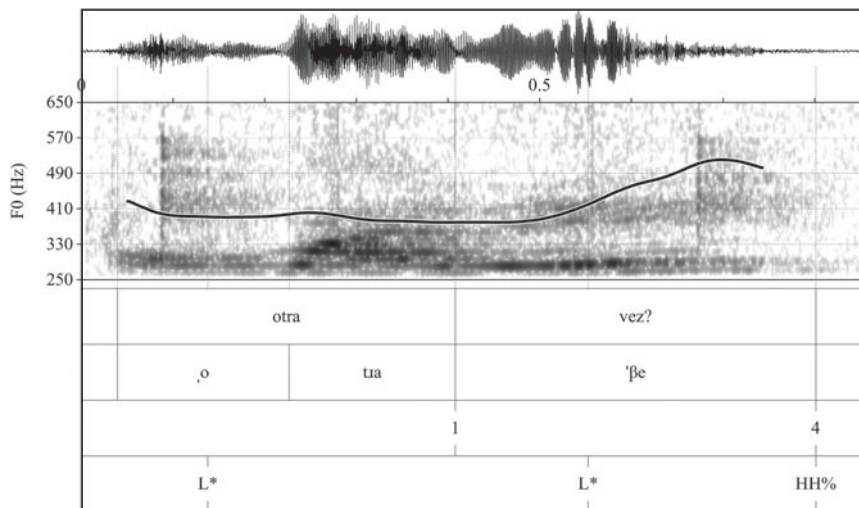


Fig. 9. Waveform display, spectrogram, fo contour and prosodic labeling of the sequence *otra vez(z)?* ‘again?’ uttered by Irene at 1;6.16.

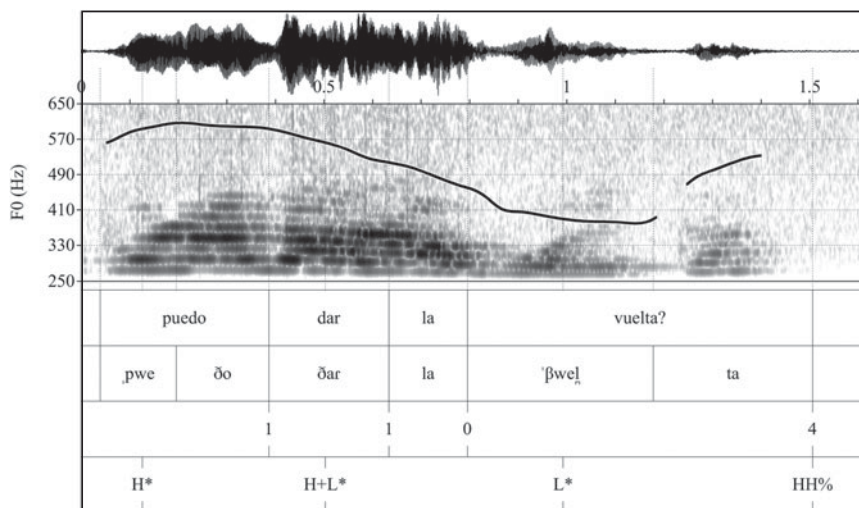


Fig. 10. Waveform display, spectrogram, fo contour and prosodic labeling of the utterance *puedo dar la vuelta?* ‘can I turn around?’ uttered by Irene at 1;11.13.

adult-like way, and they also use a variety of tunes to express requests, discontent or insistence. Importantly, the children master the tune–text alignment patterns in these contours. Gisel·la and Laura differ from the

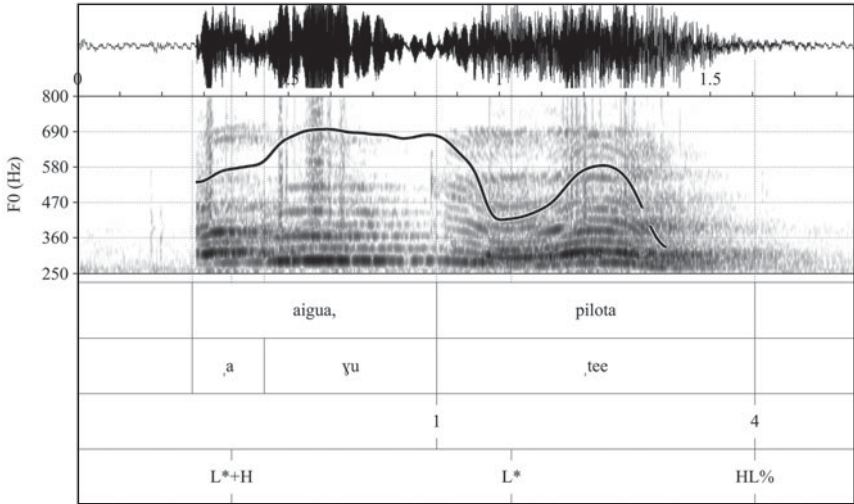


Fig. 11. Waveform display, spectrogram, fo contour and prosodic labeling of the utterance *aigua, pilota* 'water and a ball' uttered by Gisel·la at 1;10·07.

former three children in that they already show interrogative contours or the disapproval contour in the one-word period.

Figure 11 shows the first complex contour produced by Gisel·la at 1;10. The contour in this figure was produced by Gisel·la in the following context: she and her mother were reading a book, and her mother asked her a number of times what was depicted on a particular page. After answering three times, Gisel·la angrily repeated one more time to her mother. Crucially, the same contour was produced by Pep two months earlier, at 1;8, in spite of the difference in grammatical development between the two children (see Figure 7).

Figure 12 shows an interrogative utterance produced by Gisel·la at 1;7, realized as an L* nuclear contour followed by a HH% boundary tone.

In conclusion, Laura's and Gisel·la's examples of intonational development between 1;7 and 1;11 show a good phonetic and phonological command of a variety of pitch accents and boundary tones, producing them even at the one-word stage. No obvious increase in intonational grammar was attested when they started producing two-word combinations. In order to test these observations, a quantitative analysis will be presented in the next section.

Quantitative results

Intonational development. In this section, we focus on the quantitative analysis of the total number of unique NUCLEAR PITCH ACCENT

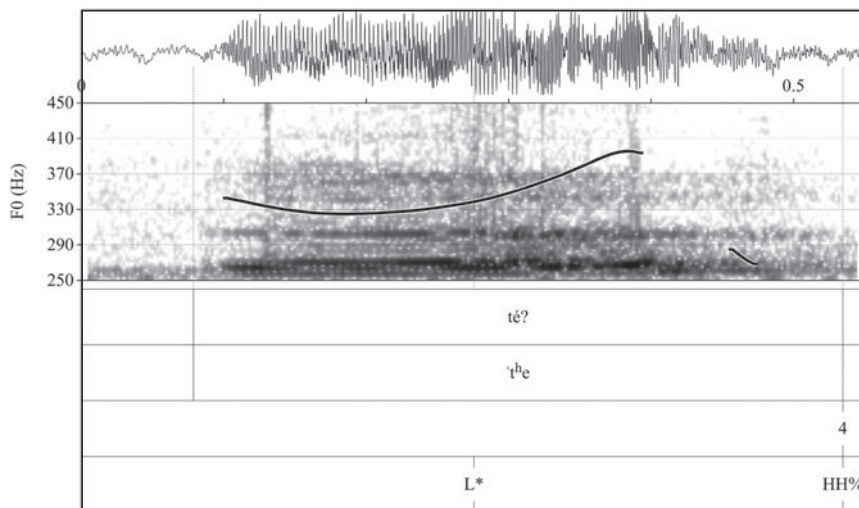


Fig. 12. Waveform display, spectrogram, fo contour and prosodic labeling of the utterance *té?* 'do you want it?' uttered by Gisel·la at 1;7·10.

CONFIGURATIONS produced by the children in each session, in other words the intonational 'lexicon' used in each session. As is well known, the nuclear pitch accent configuration is the most important part of an intonation contour; it is generally located at the end of the utterance and it is perceived as the most prominent. If an utterance has only one pitch accent, it will automatically get the nuclear pitch accent configuration. In this article, this index will be very useful because it will allow for detailed and reliable comparisons between intonational development and lexical and grammatical development.

The six stacked bar graphs in Figure 13 represent the number of different nuclear configuration types for each session of each child. Each session analyzed is represented along the x-axis; the y-axis is the number of different nuclear pitch accent configurations. The Catalan-speaking children (Pep, Guillem, Laura and Gisel·la) appear on top and the Spanish-speaking children (Irene and María) on the bottom. The graphs clearly show that: (1) all infants produce two or three distinctive nuclear pitch configurations from the onset of speech; and (2) all infants experience a 'jump', or increase in different nuclear configuration types, over the course of intonational development. Generally, the jump is located where the number of unique types increases from one or two nuclear configurations to six or seven configurations. In our view, this remarkable increase in 'intonational types', which varies in its arrival time, is equatable with

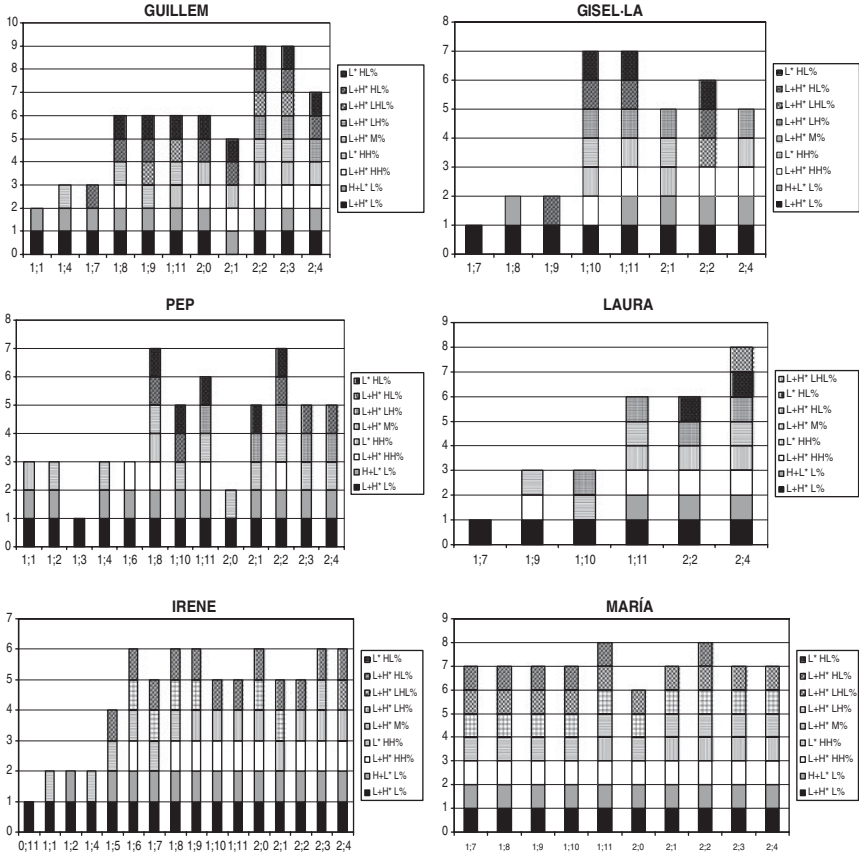


Fig. 13. Stacked bar graphs showing the number of distinctive intonation contours produced at each session, for each child. The four Catalan-speaking children (Pep, Guillem, Laura and Gisel-la) are on top and the two Spanish-speaking children (Irene and María) are on the bottom. Each session analyzed is represented along the x-axis; the y-axis is the number of different nuclear pitch accent configurations.

the first milestone event in the intonational development. Each child experiences this boost in intonational types at a given age. For Catalan, Pep and Guillem experience this shift at 1;8, while Laura and Gisel-la are at 1;11 and 1;10, respectively. For Spanish, the increase of two types arrives quite early for Irene at 1;5. She has an intonation jump from two to four types at 1;5, with an additional two more types at 1;6, meaning that she spans this increase from two to six intonation types over just two sessions. And, as noted before, María starts her dataset when she already produces eight different types of nuclear pitch accent configurations.

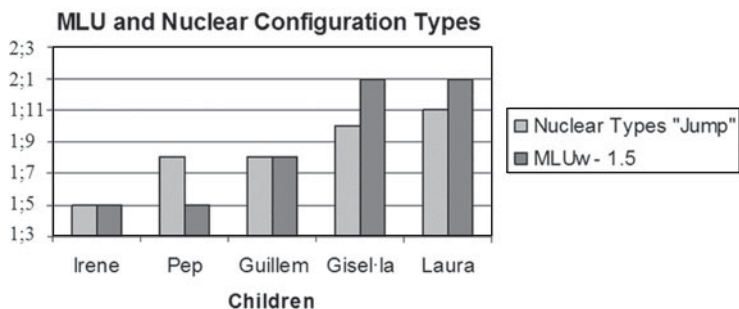


Fig. 14. Bar graph showing the age at which each child demonstrates an increase or 'jump' in the number of nuclear configuration types (light gray bar) against the age at which each child reaches an MLUw of 1.5 (dark gray bar).

Correlation between grammatical and intonational development. After obtaining these figures on nuclear configuration types, we proceed to compare the age at which the child acquires five or six different types of nuclear configurations with the age at which the child reaches an MLUw of 1.5 (estimated onset of the two-word period). Figure 14 shows a bar graph comparing the age at which each child demonstrates an increase or 'jump' in the number of nuclear configuration types (light gray bar) and the age at which each child reaches an MLUw of 1.5 (dark gray bar). The child María was not included in the graph, as there was not enough data to test the grammatical and intonational development. The comparison reveals that even though two of the children show a temporal correlation between grammatical and intonational development, the others show a delay or speed up in intonational acquisition that spans from two to four months. Two of the infants display the turning points in grammatical and intonational development during the same month, Irene at 1;5 and Guillem at 1;7. As for Pep, he reaches an MLUw of 1.5 three months before his jump in nuclear configuration types. All three of these children have a relatively early onset of the two-word period. In comparison, Gisel·la and Laura show their boost in intonational development several months before they reach an MLUw of 1.5. The graph also illustrates that Gisel·la and Laura have a slight delay in intonational and grammatical development. Although reaching the milestones later, the graph shows that they have an important understanding of intonational grammar by 1;10 and 1;11, well before they reach the two-word stage (2;1).

Thus, as is clear from Figure 14, there is no necessary temporal correlation between grammatical development (i.e. the start of the two-word period) and intonational development (i.e. the production of a variety of nuclear pitch accent configurations). In general, intonational development, with the exception of Pep, precedes grammatical development. Similarly, in Frota and

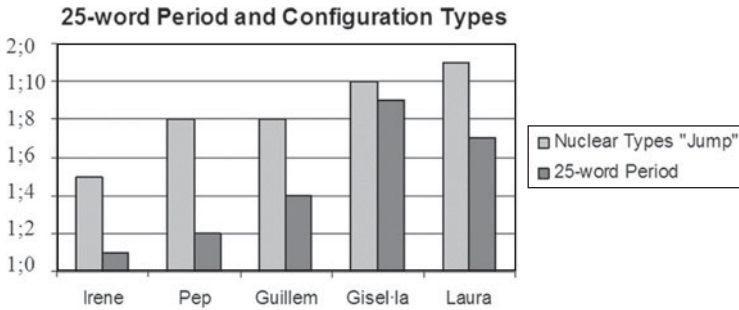


Fig. 15. Bar graph showing the age at which each child demonstrates an increase or 'jump' in the number of nuclear configuration types (light gray bar) against the age at which each child reaches a vocabulary size of 25 words (dark gray bar).

Vigário's (2008) study, the jump (i.e. the consistent use of five or more contours) occurs at 1;5, whereas the 1.5 MLU appears at 2;2.

Correlation between lexical and intonational development. As mentioned before, some investigations have reported that infants begin to use adult-like intonation contours at the 20- or 25-word point (see Vihman & DePaolis, 1998; Vihman *et al.*, 1998, for English and French; Frota & Vigário, 2008, for Portuguese). Figure 15 shows a bar graph comparing the age at which each child demonstrates the increase or 'jump' in the number of nuclear configuration types (light gray bar) and the age at which each child reaches a vocabulary size of 25 words (dark gray bar). Again María was not included in this graph because her data provide no test of the relationship between lexical and intonational development. In general, the data shows that intonational development is temporally 'linked' to lexical knowledge, as for all children the 25-word point appears before the intonational boost. The data also show that children show a closer temporal correlation between the lexical and intonational milestones, and that all of the children have this intonational acquisition after the 25-word point. While Irene and Guillem attain the 25-word period four months before the intonational boost, other children like Laura have the intonational boost one month after the 25-word point.

All in all the data corroborate previous findings that children may require some lexical knowledge (at least 25 words) to be able to show an increase in intonational development (see DePaolis *et al.*, 2008, for a review).

DISCUSSION

The development of intonational grammar

One of the goals of this article was to analyze over time the patterns of intonational development from four Catalan-speaking children and two

Spanish-speaking children. The data analyzed consist of a spontaneous corpus of 6558 meaningful utterances. One of the findings of this study has been that Catalan and Spanish children displayed an early appropriate use of distinct tunes for specific pragmatic meanings. The analysis of the data has shown that the six Catalan and Spanish children mastered the production of a wide variety of language-specific nuclear tonal configurations within an age range of 1;3 and 1;11. The results also show evidence that infants use a variety of *fo* intonation patterns to signal communicative intent, also confirming earlier accounts that the use of intonation for conveying the same meanings expressed by the adult language is present from the onset of speech (Cruttenden, 1982; Marcos, 1987; Thorson, Borràs-Comes, Crespo-Sendra, Vanrell & Prieto, 2009). In a study of ten infants acquiring French, Marcos (1987) found that rising *fo* patterns were used more frequently in both initial requests and repeated requests than in labeling activities. Similarly, Thorson *et al.* (2009) investigated in detail *yes/no* interrogative forms produced by the Catalan- and Spanish-acquiring group of children investigated here between the ages of 1;0 and 2;4, for a total of 733 interrogatives. Importantly, the data show that the variety of *yes/no* questions produced by the children do in fact reflect the adult inventory of intonational patterns, which were previously investigated in the child-directed speech data. Importantly, the associated pragmatic meaning was also adult-like from the beginning of the children's productions.

Recent cross-linguistic evidence on the early production of language-specific pitch contours backs up the results from Catalan and Spanish. For European Portuguese, Frota and Vígário (2008) have reported that a European Portuguese child acquired the inventory of pitch accents and boundary tones in an adult-like way at 1;9, with the emergence of such contours as early as 1;5. Recently, Chen and Kent (2009) have analyzed the prosodic patterns produced by Mandarin-learning infants at the onset of speech. They report that the distribution *fo* patterns showed significant similarities in babbling and early words, and that these distributions were also similar to their caregivers' data. This cross-linguistic evidence seems to suggest that *fo* alignment patterns are produced quite robustly in early production. Indeed, in our study, fine control of tune-text alignment was also described for all meaningful productions, and consequently no stress errors were reported in the data. The Catalan and Spanish data has shown that children master the tune-text alignment of the target intonation contours from the production of their first words. By contrast, it is only over the course of several months that they improve upon the scaling of sentence-final low boundary tones. Corroborating evidence for the early control of *fo* alignment and tune-text association comes from a variety of studies. For example, Astruc *et al.* (2009) analyzed naming data from twenty-four two-, four- and six-year-old English, Spanish and Catalan

children and showed that in rising accents of the type L+H* L% that children as young as two control relevant intonation parameters such as pitch height and pitch timing, although they still do not control syllabic duration and they still lengthen excessively word-final syllables. Kehoe *et al.* (1995) also found that English infants aged 1;6 controlled the implementation of *f₀*, intensity and duration patterns to indicate stress in elicited trochaic words. Vihman and DePaolis (1998) and Vihman *et al.* (1998) showed that English and French infants at the 25-word point are able to produce adult-like *f₀* patterns to mark stress. Finally, for European Portuguese, Frota and Vigário (2008) showed that while the precise alignment of the leading nuclear tone in H+L* pitch accents in statements is not adult-like until 1;9, the alignment of the L+H* pitch accent is adult-like after 1;2.

The early *f₀* control in the production of intonation patterns should not come as a surprise, given that perception studies in newborns and babies have repeatedly shown that babies are extremely sensitive to the prosody of their native languages. Infants have been shown to be sensitive to the predominant stress patterns of their languages (see Jusczyk, Cutler & Redanz, 1993, for English), something that helps them to start acquiring the lexicon and syntax of their native language (Christophe *et al.*, 1997; Christophe *et al.*, 2003; Nespor *et al.*, 1996; among many others). Thus, given this substantial capability in the processing of prosodic information, we can expect that these prosodic patterns will be reflected in infant babble and early productions. Not surprisingly, the control of pitch in imitation has been documented in infants as early as 0;3 (Papoušek & Papoušek, 1989).

Yet the literature on the acoustic and prosodic characteristics of babbling is partially contradictory and it is not clear yet how early infants' vocalizations are influenced by the adult prosodic system. Even though there are some studies that do not detect language-specific differences in the babble of infants aged 1;0 or 1;6 (see for example Engstrand, Williams & Lacerda, 2003), others have reported that some children use adult-like intonation in the late babbling period (Crystal, 1986; Chen & Kent, 2009; Dore, 1975; see Snow & Balog, 2002, for a review), a phenomenon described as 'jargon intonation' or 'the tune before the words'. The idea that the emergence of intonation patterns is related to the onset of speech is consistent with a number of diary studies and other investigations indicating that children begin to use one or more contours at about 1;0 or 1;1 (Crystal, 1986; Halliday, 1975). Yet different reports in the literature show that there is no clear consensus as to whether intonation in the majority of children develops early (with respect to the onset of speech) or relatively late. DePaolis *et al.* (2008: 408) conclude that: "Taking all of these studies together, there appears to be limited evidence for the control of *f₀* in the

pre-linguistic period but a clear consensus that, by the time of regular production of multiword combinations, *f₀* has become decidedly adult-like.”

In our view, some of the discrepant results in the literature may be due to the fact that investigations have analyzed the patterns of fundamental frequency, duration and intensity together in the infant’s production, not taking into consideration potentially different developmental patterns of individual parameters (see DePaolis *et al.*, 2008; among many others). While there is evidence that infants are able to control some of the *f₀* characteristics at an early age, other prosodic correlates, such as timing or intensity patterns, are probably acquired later, giving a potential erroneous picture on the early prosodic patterns produced by the children (for a review, see DePaolis *et al.*, 2008).

Even though the children in our study finely controlled the *f₀* alignment patterns in their early productions, they did not produce other acoustic parameters like the duration patterns or tonal scaling in a target-like way. As in previous studies, it was clear that the timing patterns, as segmental patterns, were not target-like from the earliest productions and developed more slowly than intonation patterns. For example, Kehoe and collaborators tested English children from 1;8 to 3;0 and found that only the older children produced appropriate stressed–unstressed durational contrasts (Kehoe *et al.*, 1995; Kehoe & Stoel-Gammon, 1997). Snow (1994) showed that children started to control final lengthening after the onset of the multiword stage (1;5–2;0), but they experienced a regression a few months later (see also Snow, 2006). In Frota and Matos (2008), the same child analyzed in Frota and Vigário (2008) was observed for duration patterns. It was shown that final lengthening was not produced at 1;9, but was already in place at 2;2, at the onset of the two-word stage.

Other phonetic implementation discrepancies with the adult language productions were found with respect to the control of tonal scaling. For example, the target low boundary tones (L%) in statements were frequently not fully produced. In those cases, the L% boundary tone was realized as a mid tone by the child, and not as the target low tone found in adult speech. Although the target level was not accomplished, the prosodic meaning of the utterance was retained. Previous investigations have also pointed out the lack of control of pitch range and tonal scaling in infants’ early productions (Astruc *et al.*, 2009; Vanrell *et al.*, 2010; Lleó *et al.*, 2004; Lleó & Rakow, 2011; for a review, see Snow & Balog, 2002: 1035).

From a methodological point of view, this study has shown that the Autosegmental Metrical framework can be successfully applied to investigations of early intonational development (see also Prieto & Vanrell, 2007, for Catalan; Chen & Fikkert, 2007, for Dutch; Frota & Vigário, 2008, for European Portuguese; Thorson *et al.*, 2009, for Catalan and Spanish).

Data from the four languages (Catalan, Dutch, European Portuguese and Spanish) indicate that children produce target-like intonation patterns from the beginning of their productions and thus they can be successfully analyzed in terms of pitch accents and boundary tones. In our view, the use of this model to analyze prosodic development provides us with a strong tool for analyzing intonation patterns in terms of phonologically distinct contours. An AM-based analysis will allow for more detailed studies on the phonetic implementation of pitch alignment and scaling in those contours. As pointed out by Chen and Fikkert (2007), even though the contour-based approach has proven useful for describing early intonation of early babbling, it falls short when trying to describe the early intonation patterns found in late babbling and early speech.

The biological hypothesis

The findings from this study also have implications for the widely held idea that early intonational productions might reflect biological and physiological universals. The fact that many studies on child language production data find that the falling contour is predominant over the rising contour (Behrens & Gut, 2005; Snow, 2006) has been generally attributed to a universal production mechanism, as stated in Lieberman's breath group theory (Lieberman, 1967), where a fall is the natural result of a decrease in the subglottal air pressure towards the end of a breath group. Thus falling contours were conceived to be more natural and less 'marked' than rising contours. In Snow's (2006) review of research on intonational development, he concludes that: "the precocious expression of intonation in the youngest infants pointed to the role of physiological universals and emotional experience. It is concluded that children's early intonation reflects biological, affective, and linguistic influences." This explanation has even been held to explain the productions of falling contours in two-word utterances. For example, in a case study on the prosodic and syntactic organization of a German-acquiring child's two-word utterances, Behrens and Gut (2005) analyzed the intonation of the child's two-word utterances produced over a period of three months. They observed that the falling contours were most frequent across all types of utterances and that rising contours were rarely used.

There are several arguments that call into question the physiologically based explanation in early speech. First, prior work on intonational development has focused only on the analysis of overall contour shape. This method basically classified pitch contours into two possible patterns, falling contours and rising contours. Yet recent work on the development of intonational patterns in Dutch, European Portuguese and Catalan, and now Spanish, have show that children produce more complex patterns of nuclear

pitch configurations from the onset of speech, thus indicating that the classification of contours into rising and falling contours represents an oversimplification of the data that does not allow us to discover whether the children are using more complex *f₀* patterns.

Second, it is also clear that in Romance and Germanic languages the predominant *f₀* contour in adult speech and in child-directed speech is the falling contour, which is the typical intonational form of statements. Falling contours are far more common than rising contours, which tend to encode interrogative and continuation meanings. It is thus not surprising that children tend to produce those contours more frequently in their speech. As for the production of interrogative forms, especially telling is the case of Catalan, which has both falling and rising intonations for informational *yes/no* questions. In a study of the acquisition of those patterns by four Catalan-speaking infants (Thorson *et al.*, 2009), they always produced the rising pattern before the falling one. For example, Gisel·la produced 96 instances of the rising *yes/no* questions and just one falling *yes/no* question between the ages of 1;10 and 2;1, the period in which she starts producing the interrogative forms. Laura, on the other hand, produced 72 rising *yes/no* questions and one falling *yes/no* question between 1;9 and 2;2. Finally, Guillem produced 96 rising interrogative questions and 26 falling questions in just one of the first sessions where he begins using interrogatives. It is also important to note that the most frequent patterns of interrogatives in child-directed speech were the rising patterns (that is, L+H* HH%, and after L* HH%).

Finally, it is also clear that the first intonational contours produced by the Catalan and Spanish infants under study contain a rising pitch accent (L+H*) associated with the nuclear stressed syllable, a clear indication that children are able to finely control *f₀* movements from the onset of speech. Thus, the fact that the majority of intonational contours corresponding to statements are falling should not be taken as a straight argument in favor of the physiological tendency to lower the fundamental frequency in the course of a sentence. Following this view, it is rather surprising that early productions reveal that infants undershoot the low target *f₀* values at the end of the sentence.

Relationship between lexical, grammatical and intonational development

One of the overarching goals of this article was to investigate whether prosody drives syntactic and lexical development in early production. The grammatical complexity measure used is the Mean Length of Utterance in words (MLUw). Lexicon or vocabulary size was computed with the 'freq' command in CLAN for CHILDES by listing the number of unique recorded words produced by each child per session. Finally, a

measure of the ‘intonational lexicon’ was computed by analyzing the number of distinctive nuclear pitch accent configurations produced in each session. These indices have been proven to be very useful, as they allow for quantitative comparisons between intonational, lexical and grammatical development.

The quantitative analyses of the data presented earlier demonstrate the following: (a) all Catalan- and Spanish-speaking infants produce a handful of target-like nuclear pitch accent configurations from the onset of speech (see Figure 13) – these configurations are typically statements (L+H* L%, H+L* L%), focal statements, and vocatives of different types (L+H* !H%, L+H* L!H%); (b) Catalan and Spanish infants experience a ‘jump’, or increase in different nuclear configuration types, over the course of intonational development – this is the time where children use six to seven types of tunes in a consistent way; (c) there is no clear temporal relationship between the start of the two-word period and the ‘jump’ in the number of distinctive nuclear configuration types. Even though two of the children show a temporal coincidence between grammatical and intonational developments (Irene and Guillem), two other children (Gisel·la and Laura) acquire intonation before the two-word period. It is also possible to show a delay of intonational development with respect to the start of the two-word period – cf. Figure 14; (d) finally, there is no clear temporal relationship between the age at which the children reach a vocabulary size of 25 words and the first establishment of intonational grammar – cf. Figure 15. Yet an important generalization is that all the children show this intonational burst after the 25-word point (between one and six months later, depending on the child).

A close relationship between the presence of a small lexicon (20- or 25-word vocabulary) and an increase in intonational development has been mentioned by previous studies (see Vihman & DePaolis, 1998; Vihman *et al.*, 1998, for English and French; Frota & Vigário, 2008, for Portuguese). As DePaolis *et al.* (2008: 417) point out at the end of their article: “more finely tuned use of prosody may require a level of attention to linguistic detail that begins to be possible only as word production becomes well established.”

In our data, we can argue that the intonation jump always follows the 25-word point and generally precedes the two-word stage (yet see Pep, who represents the only exception). In our view, the relative independence between the start of more complex structures and intonation can be traced back to the temporal independence between lexical and syntactic developments. As we can observe by comparing the graphs in Figures 2 and 3, while Gisel·la and Laura get to the two-word stage when they have an approximate vocabulary size of 100 words, Guillem (and to a certain extent, Pep) gets to the two-word stage quite early, at 1;8 (five months before

Laura and Gisel·la), while he does not attain a lexicon size of 100 words until he is 2;4. This clearly suggests that vocabulary size and grammatical complexity measures are not strictly correlated in development.

CONCLUSION

This article examines developmental data from four Catalan-speaking children and two Spanish-speaking children between the ages of approximately 1;0 and 2;4. A total number of 6558 meaningful utterances were analyzed prosodically and assessed for their pragmatic meaning. In the analysis, we focused on the relationship between lexical and grammatical development and the development of intonational grammar (that is, the capacity to use appropriate intonation for specific pragmatic meanings).

The results indicate that the six Catalan and Spanish children produce the basic phonologically distinct *f₀* contours of their ambient language from the onset of their speech. A few months later, each child exhibits a ‘jump’ in the number of nuclear configuration types, varying only at what age this increase occurs, thus showing an important knowledge of the adult intonational grammar. Importantly, our data show evidence that infants use these *f₀* patterns in a pragmatically adequate way to signal communicative intent, also confirming some earlier accounts (see also Cruttenden, 1982; Marcos, 1987; Thorson *et al.*, 2009). Recent data from two other languages (Dutch and European Portuguese) also find that children have largely acquired the adult inventory of pitch accents and boundary tones before the age of two (Chen & Fikkert, 2007, for Dutch; Frota & Vigário, 2008, for European Portuguese). It is worth noting that other languages are different with regard to tune–text alignment, as in the case of falling accents in European Portuguese and Dutch child speech, and that this fact might also be influencing early intonational development.

The Catalan and Spanish data at hand show that children master the tune–text alignment of a handful of pitch accents and boundary tones from the onset of speech, and it is over the course of several months that they improve upon the scaling of low boundary tones. Corroborating evidence for the early control of *f₀* alignment and association comes from a variety of studies (Astruc *et al.*, 2009; Kehoe *et al.*, 1995; Vihman & DePaolis, 1998; Vanrell *et al.*, 2010; Vihman *et al.*, 1998).

From a methodological point of view, this study demonstrates that the Autosegmental Metrical model of intonation, and specifically the inventory of adult Spanish and Catalan pitch accents and boundary tone combinations (Cat_ToBI and Sp_ToBI: Prieto *et al.*, 2009; Prieto, in press; Estebas-Vilaplana & Prieto, 2010) can be successfully applied to the analysis of early intonation patterns produced by Catalan and Spanish infants. In

our view, the application of this model to the analysis of early fo patterns cross-linguistically can represent an important tool that will allow us to evaluate both the form and functions of early intonation patterns in relation to the target patterns.

Some important conclusions of this study are related to the potential temporal correlations between lexical and intonational development and between grammatical and intonational development. First, our results demonstrate that, contrary to what has been claimed in the literature, children's emerging intonation is not correlated in time with grammatical development. While some children reach the grammatical and intonational milestones at the same time (Irene and Guillem), others display the intonational burst several months after the two-word period began (Pep), and others (Gisella and Laura) show an important knowledge of intonational grammar well before they produce two-word combinations. Second, our study suggests a relatively close temporal correlation between lexical development and intonational development, in the following sense. First, all children are able to produce a handful of intonation contours from the production of their first words. Second, all children display a burst in intonational production after they acquired a critical mass of words, namely, 25 lexical items. Studies by Frota and Vigário (2008), Vihman and DePaolis (1998), Vihman *et al.* (1998) and DePaolis *et al.* (2008), among others, support the idea that prosodic competence requires some lexical knowledge. More research is needed to evaluate whether there is a more precise correlation between the number of lexical words acquired and the child's prosodic development.

Taken together, these results seem to indicate that the emergence of the intonational grammar of the ambient language is closely related in time with the onset of speech. We need to further investigate whether these intonation patterns systematically reflect target pragmatic meanings or whether there is any interaction between the acquisition of target intonation patterns and their semantic function (i.e. in the case of interrogative sentences). Another pending question is whether late babbling patterns, produced in the same period of time, also support the hypothesis of continuation and reflect adult-like intonational patterns, as some recent studies seem to suggest (see Chen & Kent, 2009; DePaolis *et al.*, 2008; Esteve-Gibert, 2010, among others).

REFERENCES

- Aguilar, L., de-la-Mota, C. & Prieto, P. (coords) (2009a). *Cat_ToBI Training Materials*. <http://prosodia.upf.edu/cat_tobi/>.
- Aguilar, L., De-la-Mota, C. & Prieto, P. (coords) (2009b). *Sp_ToBI Training Materials*. <http://prosodia.upf.edu/sp_tobi/>.

- Astruc, L., Prieto, P., Payne, E., Post, B. & Vanrell, M. M. (2009). Acquisition of tonal targets in Catalan, Spanish, and English. In A. Appleton, E. Lash & M. L. Jøhndal (eds), *Cambridge Occasional Papers in Linguistics*, Volume 5, 1–14.
- Austin, J. L. (1962). *How to do things with words*. London: Oxford University Press.
- Beckman, M. & Pierrehumbert, J. B. (1986). Intonational structure in English and Japanese. *Phonology Yearbook* 3, 255–310.
- Behrens, H. & Gut, U. (2005). The relationship between prosodic and syntactic organization in early multiword speech. *Journal of Child Language* 32, 1–34.
- Boersma, P. & Weenink, D. (2009). Praat: Doing phonetics by computer (Version 5.1.12) [Computer program]. Retrieved 4 August 2009, from www.praat.org/
- Chen, A. & Fikkert, P. (2007). Intonation of early two-word utterances in Dutch. In J. Trouvain and W. J. Barry (eds), *Proceedings of the XVth International Congress of Phonetic Sciences*, 315–20. Pirrot GmbH: Dudweiler.
- Chen, L. M. & Kent, R. (2009). Development of prosodic patterns in Mandarin-learning infants. *Journal of Child Language* 36, 73–84.
- Christophe, A., Gout, A., Peperkamp, S. & Morgan, J. (2003). Discovering words in the continuous speech stream: The role of prosody. *Journal of Phonetics* 31, 585–98.
- Christophe, A., Guasti, M. T., Nespor, M., Dupoux, E. & van Ooyen, B. (1997). Reflections on prosodic bootstrapping: Its role for lexical and syntactic acquisition. *Language and Cognitive Processes* 12, 585–612.
- Cruttenden, A. (1982). How long does intonation acquisition take? *Papers and Reports on Child Language Development* 21, 112–18.
- Crystal, D. (1973). Non-segmental phonology in language acquisition: A review of the issues. *Lingua* 32, 1–45.
- Crystal, D. (1986). Prosodic development. In P. J. Fletcher & M. Garman (eds), *Studies in first language development*, 174–97. New York: Cambridge University Press.
- DePaolis, R. A., Vihman, M. M. & Kunnari, S. (2008). Prosody in production at the onset of word use: A cross-linguistic study. *Journal of Phonetics* 36, 406–422.
- D’Odorico, L. & Carubbi, S. (2003). Prosodic characteristics of early multi-word utterances in Italian Children. *First Language* 23(1), 97–116.
- D’Odorico, L. & Fasolo, M. (2009). The prosody of early multi-word speech: Word order and its intonational realization in the speech production of Italian children. *Enfance* 61(3), 317–27.
- Dore, J. (1973). The development of speech acts. Unpublished doctoral dissertation, City University of New York.
- Dore, J. (1974). A pragmatic description of early language development. *Journal of Psycholinguistic Research* 4, 343–50.
- Dore, J. (1975). Holophrases, speech acts and language universals. *Journal of Child Language* 2, 21–40.
- Engstrand, O., Williams, K. & Lacerda, F. (2003). Does babbling sound native? Listener responses to vocalizations produced by Swedish and American 12- and 18-month-olds. *Phonetica* 60, 17–44.
- Estebas-Vilaplana, E. & Prieto, P. (2010). Peninsular Spanish intonation. In P. Prieto & P. Roseano (coords). *Transcription of intonation of the Spanish language*, 17–48. München: Lincom Europa.
- Esteve-Gibert, N. (2010). The development of prosodic patterns in Catalan-babbling infants. Unpublished MA thesis, Universitat Pompeu Fabra.
- Frota, S. & Matos, N. (2008). O tempo no tempo: um estudo do desenvolvimento das durações a partir das primeiras palavras. In Alexandra Fiéis & M. Antónia Coutinho (eds), *Textos Seleccionados do XXIV Encontro Nacional da Associação Portuguesa de Linguística*, 281–95. Lisboa: Colibri/APL.
- Frota, S. & Vigário, M. (2008). The intonation of one-word and first two-word utterances in European Portuguese. Paper presented at the *Third Conference on Tone and Intonation* (TIE 3), Lisbon, 15–17 September 2008.

- Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- Halliday, M. A. K. (1975). *Learning how to mean: Explorations in the development of language*. London: Edward Arnold.
- Jun, S. A. (ed.) (2005). *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press.
- Jusczyk, P. W., Cutler, A. & Redanz, N. J. (1993). Infants' preference for the predominant stress pattern of English words. *Child Development* **64**, 675–87.
- Kehoe, M. & Stoel-Gammon, C. (1997). Truncation patterns in English-speaking children's word productions. *Journal of Speech, Language, and Hearing Research* **40**, 526–41.
- Kehoe, M., Stoel-Gammon, C. & Buder, E. H. (1995). Acoustic correlates of stress in young children's speech. *Journal of Speech and Hearing Research* **38**, 338–50.
- Ladd, D. R. (2008 [1996]). *Intonational phonology*, 2nd edn. Cambridge: Cambridge University Press.
- Lieberman, P. (1967). *Intonation, perception, and language*. Cambridge, MA: MIT Press.
- Lleó, C. & Rakow, M. (2011). Intonation targets of yes-no questions by Spanish and German monolingual and bilingual 2;0- and 3;0-year-olds. In T. Kupisch & E. Rinke (eds), *The development of grammar: Language acquisition and diachronic change*, 213–34. Hamburger Studies on Multilingualism **11**. Amsterdam; Philadelphia: John Benjamins.
- Lleó, C., Rakow, M. & Kehoe, M. (2004). Acquisition of language-specific pitch accent by Spanish and German monolingual and bilingual children. In T. L. Face (ed.), *Laboratory approaches to Spanish phonology*, 3–27. Berlin; New York: Mouton de Gruyter.
- MacWhinney, B. & Snow, C. (1985). The Child Language Data Exchange System. *Journal of Child Language* **12**, 271–96.
- Marcos, H. (1987). Communicative functions of pitch range and pitch direction in infants. *Journal of Child Language* **14**, 255–68.
- Nespor, M., Guasti, M. T. & Christophe, A. (1996). Selecting word order: The Rhythmic Activation Principle. In U. Kleinhenz (ed.), *Interfaces in Phonology*, 1–26. Berlin: Akademie Verlag.
- Ninio, A. (1992). The social bases of Cognitive/Functional Grammar: Commentary on Tomasello, M. (1992), *The social bases of language acquisition*. *Social Development* **1**, 155–58.
- Ninio, A., Snow, C. E., Pan, B. A. & Rollins, P. R. (1994). Classifying communicative acts in children's interactions. *Journal of Communication Disorders* **27**, 158–87.
- Papoušek, M. & Papoušek, H. (1989). Forms and functions of vocal matching in interactions between mothers and their precanonical infants. *First Language* **9**, 137–58.
- Pierrehumbert, J. B. (1980). The phonology and phonetics of English intonation. Unpublished PhD dissertation, MIT.
- Prieto, P. (2006). The relevance of metrical information in early prosodic word acquisition: A comparison of Catalan and Spanish. *Language and Speech* (special issue on the Crosslinguistic Perspectives on the Development of Prosodic Words, ed. by K. Demuth), **49**(2), 233–61.
- Prieto, P. (in press). The intonational phonology of Catalan. In S. A. Jun (ed.), *Prosodic typology 2*. Oxford: Oxford University Press.
- Prieto, P., Aguilar, L., Mascaró, I., Torres-Tamarit, F. J. & Vanrell, M. M. (2009). L'etiquetatge prosòdic Cat_ToBI. *Estudios de Fonética Experimental* **XVIII**, 287–309.
- Prieto, P. & Vanrell, M. M. (2007). Early intonational development in Catalan. In J. Trouvain & W. J. Barry (eds), *Proceedings of the XVth International Congress of Phonetic Sciences*, 309–314. Dudweiler: Pirrot GmbH.
- Randolph, J. J. (2008). *Online Kappa Calculator*. Retrieved 10 March 2010, from <http://justus.randolph.name/kappa>
- Rose, Y., MacWhinney, B., Byrne, R., Hedlund, G., Maddocks, K., O'Brien, P. & Wareham, T. (2006). Introducing Phon: A software solution for the study of phonological acquisition. In D. Bammann, T. Magnitskaia & C. Zaller (eds), *Proceedings of the 30th*

- Annual Boston University Conference on Language Development*, 489–500. Somerville, MA: Cascadilla Press.
- Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. London: Cambridge University Press.
- Snow, C. (1994). Beginning from baby talk: Twenty years of research on input and interaction. In C. Gallaway & B. Richards (eds), *Input and interaction in language acquisition*, 3–12. Cambridge: Cambridge University Press.
- Snow, D. (2000). The emotional basis of linguistic and nonlinguistic intonation: Implications for hemispheric specialization. *Developmental Neuropsychology* **17**, 1–28.
- Snow, D. (2006). Regression and reorganization of intonation between 6 and 23 months. *Child Development* **77**, 281–96.
- Snow, D. & Balog, H. L. (2002). Do children produce the melody before the words? A review of developmental intonation research. *Lingua* **112**, 1025–58.
- Thorson, J., Borràs-Comes, J., Crespo-Sendra, V., Vanrell, M. M. & Prieto, P. (2009). The acquisition of melodic form and meaning by Catalan and Spanish speaking children. Paper presented at Phonetics and Phonology in Iberia 2009, Las Palmas de Gran Canaria, Spain.
- Vanrell, M. M., Prieto, P., Astruc, Ll., Payne, E. & B. Post (2010). Early acquisition of F₀ alignment and scaling patterns in Catalan and Spanish. *Speech Prosody 2010* 100839:1–4, <http://speechprosody2010.illinois.edu/papers/100839.pdf> (last accessed 26 June 2010).
- Vihman, M. M. & DePaolis, R. A. (1998). Perception and production in early vocal development: Evidence from the acquisition of accent. In M. C. Gruber, D. Higgins, K. S. Olson & T. Wysocki (eds), *Chicago Linguistic Society 34*, Part 2: Papers from the panels, 373–86. Chicago, IL: CLS.
- Vihman, M. M., DePaolis, R. A. & Davis, B. L. (1998). Is there a ‘trochaic bias’ in early word learning? Evidence from infant production in English and French. *Child Development* **69**, 935–49.
- Yoon, T., Chavarria, S., Cole, J. & Hasegawa-Johnson, M. (2004). Intertranscriber reliability of prosodic labeling on telephone conversation using ToBI. In *Proceedings of ICSA International Conference on Spoken Language Processing*, 2729–32. Jeju Island, Korea, 4–8 October 2004.
- Yoonsook, M., Cole, J. & Lee, E.-K. (2008). Prosody perception of naïve listeners: Evidence from large multi-transcribers’ reliability study. Poster presented at The 82nd Annual Meeting of Linguistic Society of America (LSA), Chicago, IL, 2–5 January 2008.