Early development of intonation: Perception and production

Chapter · January 2018
DOI: 10.1075/tilar.23.08fro

2 authors:

Sonia Frota
University of Lisbon
136 PUBLICATIONS 1,120 CITATIONS

Joseph Butler
University of Lisbon
24 PUBLICATIONS 258 CITATIONS

Some of the authors of this publication are also working on these related projects:

Eyes and Brain: Early markers of language development View project

FreP - Frequency Patterns of Phonological Objects in Portuguese: Research and Applications View project

All content following this page was uploaded by Sonia Frota on 08 February 2018.
The user has requested enhancement of the downloaded file.
CHAPTER 8

Early development of intonation

Perception and production

Sónia Frota & Joseph Butler
Universidade de Lisboa / Universidade de Lisboa

This chapter focuses on early development of intonation. Together with a precocious sensitivity to prosody documented in the literature, recent research has shown that infants’ early perception of pitch-based categories is already language-specific by 4–5 months, and that their discrimination abilities differ not only according to ambient language but also as a function of pitch properties (e.g., pitch direction, or pitch alignment). On the production side, and focusing on studies within the Autosegmental-Metrical framework, findings suggest that key landmarks in intonational development precede and constrain the acquisition of other aspects of grammar (e.g., word and phrase size, and combinatorial speech). Both from a perception and production point of view, language specific effects emerge very early on in development, underlying cross-linguistic differences.

Introduction

A precocious sensitivity to prosody is by now a well-established fact in early language development. Newborn infants have been shown to be sensitive to the global rhythmic properties that distinguish between languages (Nazzi, Bertoncini, & Mehler, 1998), to word-level rhythm, that is, to different patterns of lexical stress (Sansavini, Bertoncini, & Giovanelli, 1997), to the contrast between function words, which are unstressed, and content words, which bear stress (Shi, Werker, & Mogan, 1999), to the prosodic cues, pitch included, that signal boundaries between words (Christophe, Dupoux, Bertoncini, & Mehler, 1994), and to the differences in pitch contours that result from contrasting lexical pitch accents (Nazzi, Floccia, & Bertoncini, 1998). Furthermore, newborns were found to rely more on vowels than on consonants when recognizing words, unlike adults and toddlers (Benavides-Varela, Hochmann, Macagno, Nespor, & Mehler, 2012). This is consistent with a precocious sensitivity to prosody, given that vowels are the main
carriers of prosody. Although these studies suggest that newborns’ sensitivity to prosody may not depend on language experience, they leave open the possibility that this early sensitivity may evolve differently according to the language-specific prosodic features of the ambient language. In this chapter, we focus on early development of intonation drawing on recent findings for several languages from perception and production in infants and young children.

Studies into the early perception of prosody have mostly examined the acquisition of lexical prosodic contrasts, namely word stress and lexical pitch (on word stress, Bhatara, Boll-Avetisyan, Hohle, & Nazzi, this volume). For lexical pitch, it has been shown that learners of tone languages (e.g., Mandarin), unlike learners of non-tone languages (e.g., English), show stable discrimination of distinct lexical tones throughout the first year of life, and as early as 4–5 months (Mattock & Burnham, 2006; Mattock, Molnar, Polka, & Burnham, 2008; Yeung, Chen, & Werker, 2013). Similarly, learners of pitch accent languages (e.g., Japanese) are able to discriminate native pitch accent contrasts from as early as 4 months (Sato, Sogabe, & Mazuka, 2009). Interestingly, phonetic variability seems to hinder the discrimination of native lexical tones (Shi, 2010, using variable tone contexts), whereas the perception of native pitch accent is not affected (Sato et al., 2009, using segmental variation). Much less is known about how the perception of intonation develops, and whether parallels may be found with lexical pitch, or lexical prosody in general. This topic only very recently has gained researchers’ attention. The second section of the present chapter is devoted to early perception of intonation and its developmental path. We review the recent growing body of literature, and discuss the findings as evidence for general perceptual abilities for pitch-based contrasts and/or for the early role played by language experience.

Unlike perception, studies on intonational development from the viewpoint of production have been around for nearly three decades. Most of these studies followed a contour-based approach to the development of intonation, focusing on overall shapes and acoustic measures of pitch over the whole utterance or its terminal contour (e.g., DePaolis, Vihman, & Kunnari, 2008; Esteve-Gibert & Prieto, 2013; Hallé, Boysson-Bardies, & Vihman, 1991; Frota & Vigário, 1995; Snow, 1994, 2002, 2006; see Snow & Balog, 2002, for a review). They have revealed cross-linguistic differences in the type of contours (falling or rising) first produced by the child, consistent with communicative intentions and intonational differences in the ambient languages. However, the contour-based approach is not very informative about the structural properties of intonation, such as the types of nuclear pitch accent and boundary tone, the distribution of pitch events, or the ways in which the tune aligns with the text. These properties are known to be crucial to the description of any intonational system (Gussenhoven, 2007; Ladd, 2008), and it is thus critical to understand how they are acquired. In the Autosegmental-Metrical
(AM) framework, currently the most widely used approach to intonation (e.g., Frota & Prieto, 2015a; Jun, 2014a), these properties are at the core of intonational description and analysis. Studies of early intonation within the AM framework are however rare. In Section 3 of this chapter, we focus on these studies, review the results reported, and discuss similarities and differences across languages in the path of intonational development and its relation to other aspects of grammatical development.

The AM framework allows for the distinction between a phonological and a phonetic level in intonation, and the comparison across languages at both levels. Within this model, the intonation contour is formed by a string of high and low tonal events that associate with lexically stressed syllables or with prosodic phrase edges. The former are known as pitch accents and the latter as boundary tones, and either can be single-tone (high or low, e.g., H* or L% respectively for a pitch accent and a boundary tone) or some combination of high and low (e.g., H + L* or LH%). Recent work on early child intonation has shown that this model is adequate to investigate the development of intonation in perception and production. Therefore, our review mostly focuses on work that has used the AM approach to understand the developmental path of intonation in infancy and early childhood.

Given the only recent research focus on early development of intonation, it is not surprising that our understanding of how intonation is acquired, based on evidence from both perception and production, and using a full-fledged phonological framework such as the AM model, is still fragmentary. Understanding the developmental path of intonation towards a full adult-like competence will impact on our knowledge of the acquisition of other linguistic categories and structures, as intonation might help the infant and young child to bootstrap aspects of the grammatical organization of the native language (Höhle, 2009; Jusczyk, 1997). It will also contribute towards tracing the development of communication and social interaction, and potentially signal language impairments (Filipe, Branco, Frota, Castro, & Vicente, 2014; Koegel, Koegel, Green-Hopkins, & Barnes, 2010). Both old and new methods and techniques to examine language processing in infants and toddlers can be explored to further study early intonation. These potential avenues for research, and their implications, are discussed in the final section of this chapter.

**Early perception of intonation**

It has already been established that infants show a preference for rising contours, high pitch and expanded melodies from birth, and that they are sensitive to changes in melodic contours (e.g. Papoušek, Bornstein, Nuzzo, Papoušek, & Symmes, 1990; Trehub, Bull, & Thorpe, 1984), however grammatically relevant
intonation contrasts have only recently begun to be examined. Intonation is a prosodic dimension that varies across languages and impacts upon meaning, similarly to lexical stress, lexical tone and lexical pitch accent, but, differently, intonation conveys phrasal meanings, such as sentence type distinctions and pragmatic meaning, rather than word level meanings.

Perception of native intonation

Of the few studies that have investigated early intonation perception, most have focused on perception relating to sentence type distinctions. Best, Levitt, and McRoberts (1991) were the first to address this issue. They asked whether the perception of intonation followed the general developmental change reported for the acquisition of segments and lexical tone, known as perceptual narrowing or perceptual attunement, whereby discrimination of non-native contrasts declines in the second half of the infants’ first year of life giving way to discrimination abilities attuned to the native language. To this end, they examined both native and non-native discrimination of exclamations and wh-questions by English-learning infants. They found that at 6–8 months infants were able to discriminate both native and non-native (Spanish) contrasts, however at 10–12 months they failed to discriminate either. Although these findings appear to contradict the view that discrimination abilities are refined through native language exposure, their stimuli contained both word order (inversion in wh-questions) and lexical properties (presence of wh-word), in addition to (highly variable) prosodic cues, which may have impacted on infants' discrimination performance. Similarly, Geffen, and Mintz (2011) also investigated discrimination of the statement/yes-no question contrast, using stimuli that included word order cues as well as intonation cues, utilizing a modified version of the Headturn Preference Procedure (HPP). Intonation cues were described as larger mean pitch changes for questions. They found that English-learning infants at 7 months successfully discriminated this contrast, but, as in Best et al.’s (1991) study, it is not clear whether infants attended to intonation cues or the lexical/word order cues also present. A recent study by Geffen and Mintz (2015), again using the modified version of the HPP, found that English-learning infants at 12 months were able to distinguish between declarative and interrogative sentence types when prosodic cues were removed, leaving only the lexical distribution patterns, thus suggesting a sensitivity to word order cues.

Two studies have also investigated sentence type distinctions, using stimuli that differ only in intonation cues. Soderstrom, Ko and Nevzorova (2011), using a Visual Habituation paradigm, utilized multiword, uninverted yes-no question and statement sentences, therefore neutralizing the word order cue to questions in English. Overall, they found a preference for questions but no direct evidence
of discrimination. These findings could possibly indicate a general preference for high/rising pitch (Papoušek et al., 1990; Trehub et al., 1984). However, the large age range used in this study (4.5 to 24 months) makes it difficult to draw conclusions about the developmental trajectory of perceptual abilities from this data. Additionally, the intonation used within each sentence type in their stimuli was very variable (e.g., declarative sentences ending with either flat, falling or bell shaped contours) and may have impacted on infants’ perception of the contours.

Frota, Butler, and Vigário (2014) used single prosodic word utterances differing in prosodic cues related to the statement/yes-no question contrast in a language, European Portuguese (EP), where this contrast is marked by prosodic cues (differently to English, in EP the yes-no question counterpart of Ele gosta da Maria ‘He likes Mary’ is Ele gosta da Maria? with the same word order as the statement, but not Gosta ele da Maria?, ‘Does he like Mary?’). Using a similar Visual Habituation paradigm, they found that, at both 5–6 and 8–9 months, EP-learning infants were able to discriminate the prosodic contrast, despite segmental variability. In contrast with Soderstrom et al.’s (2011) study, Frota et al. (2014) used intonation contours consistently; a falling pattern for the declarative and a falling-rising pattern for the interrogative. To our knowledge, this was the first study to demonstrate successful infant discrimination of an intonation contrast, suggesting precocious discrimination abilities for intonation in the presence of segmental variability similar to those reported for lexical pitch accent.

Following on from this study, Butler, Vigário, and Frota (2016) investigated a further prosodic contrast in EP that cues the difference between all-new information (broad focus) and the highlighting of a particular word (narrow/contrastive focus). This distinction is primarily marked in declarative utterances by the timing of the pitch fall, which is timed early with respect to the stressed syllable of the nuclear word in broad focus and late in narrow/contrastive focus. Differently to previous studies into intonation contrasts, which have mostly focused on pitch height/direction contrasts, this study addressed a pitch timing distinction. Using the Visual Habituation paradigm, Butler et al. (2016) found that EP learning infants discriminated this contrast at 12 months, but not at 7 months, deviating from previous findings of a precocious ability to perceive pitch distinctions. These findings suggest that there may be different developmental trajectories of the perception of different prosodic contrasts (namely, pitch height/direction and pitch timing), and underlines the importance of the nature of the prosodic cues signaling a given contrast in a given language.

Unlike other studies, Frota et al. (2014) and Butler et al. (2016) provide a description of the contours that characterize the materials used, following the AM model. Given the intonational account of the contrasts, we can describe the pitch height/direction sentence type distinction as manifested by the pitch event
at the right edge of the utterance, that is as a boundary tone contrast where L% in statements and LH% in questions follow the falling pitch in the nuclear word, as shown in Figure 1. The focus-related pitch timing contrast, on the other hand, is manifested by the way the peak aligns with the text in the nuclear pitch fall. The temporal coordination of the pitch fall with the stressed syllable yields a peak just before this syllable in the broad focus case (H + L*), whereas in the narrow/contrastive focus case the peak occurs on the stressed syllable (Figure 2). In either case the melody at the right edge of the utterance is low.

![Figure 1. Intonation contours of illustrative statement (left panel) and question (right panel) stimuli, with prosodic labelling and orthographic and phonetic transcription (in SAMPA), from Frota et al. (2014).](image1)

If different prosodic contrasts have different developmental trajectories, as the perception findings suggest, this has potential implications for the acquisition of linguistic categories marked with different prosodic features. The perception/
discrimination of a given contrast is a pre-requisite, which allows the acquisition of the grammar and pragmatics that the contrast conveys in a given language. In the case of EP, an advantage of pitch direction over pitch timing may suggest an advantage for the acquisition of the sentence type distinction over the focus distinction. In a language like Neapolitan Italian, however, the statement/question distinction is cued by pitch timing (the pitch rise is timed later in questions than in statements, D’Imperio, 2002). The potential impact for language acquisition of the presence or absence of early discrimination abilities related to a specific intonation contrast that cues a given linguistic distinction, within and across languages, remains to be investigated.

The precocious discrimination abilities for lexical pitch accent and some types of intonation contrasts reported in the literature are in line with infants’ sensitivity to, and ability to utilise, pitch from a very early age. For example, infants are able to use pitch, and other prosodic cues, to infer and express affective and pragmatic meaning (Esteve-Gibert & Prieto, this volume). Infants have also been shown to be able to use differences in pitch level or pitch shape as a cue to segment clauses and as a cue to prosodic phrase boundaries in order to guide word segmentation (Carvalho, Dautriche, Milotte, & Christophe, this volume; Seidl, 2007; Shukla, White, & Aslin, 2011). German 9-month-old infants have been shown to utilise pitch in order to aid segmentation, demonstrating segmentation only when high pitch aligns with the stressed syllable, and not when high pitch precedes or follows the stressed syllable (Zahner, Schönhuber, & Braun, 2015). Although infants do not necessarily always use pitch cues in an adult-like fashion, they are sensitive to, and able to use linguistically, pitch cues very early in development.

Perception of Non-native Intonation

Recent work has begun to focus on non-native perception of intonation, to attempt to shed further light on developing perceptual abilities for intonation. Sundara, Molnar, and Frota (2015) tested English- and Basque-learning infants’ ability to discriminate the statement/yes-no question contrast, using the stimuli prepared by Frota et al. (2014) and the same type of visual habituation procedure. The reason for using the EP stimuli with English-learning infants was twofold. Firstly, using stimuli that differed only in prosodic cues would allow further explanation for previous findings of a failure to discriminate statement vs. question contrast by English-learning infants (see Soderstrom et al., 2011, and section Perception of Native Intonation above). Secondly, the authors wanted to confirm that EP learning infants’ ability to discriminate the pitch contrast in this sentence type distinction, in the face of segmental variability, was the result of their language experience. Furthermore, Basque-learning infants were also tested, because in Basque, unlike in English, yes-no questions and statements can be segmentally identical and
differ only in their intonation, like in EP. In Basque, statements typically end in a fall (L%), but only questions have a high pitch immediately before the fall (HL% – Elordieta and Hualde, 2014). Sundara et al. (2015) found that English-learning infants at 4 months failed to discriminate the statement/question contrast using EP stimuli with segmental variability, however Basque-learning infants succeeded in this task. Importantly, English-learning infants also failed when the intonation contrast was presented without segmental variability, using the same phonetic sequence with a final fall or rise (i.e., the same bisyllabic word with either H + L* L% or H + L* LH%). The English-learning infants only succeeded when absence of segmental variability was combined with a more stringent habituation criterion. These findings suggest that English-learning infants have a limited ability to differentiate boundary tones, unlike EP and Basque-learning infants, and that intonation perception is already language-specific by 4–5 months.

Frota, Butler, Lu, and Vigário (2016) also investigated non-native intonation perception in EP-learning infants, again using the visual habituation paradigm. In this study, the authors utilised a segmentally varied non-native contrast, i.e., the distinction between Mandarin Chinese Tone 1 + Tone 4 and Tone 1 + Tone 2. This contrast provides overall similar contour shapes to the statement/yes-no question prosodic contrast in EP, although with a different implementation of the falling/rising patterns throughout the segmental string. The similarity to the EP native contours resides in the final pitch height and direction features, with Tone 1 + Tone 4 showing a final fall and Tone 1 + Tone 2 a final rise preceded by a fall, like in EP statements and questions, respectively. The differences from the EP contours arise from the difference in nature of tune versus tone. In EP, the falling pitch of the nuclear pitch accent (H + L*) in statements and questions is already in the first syllable, whereas in Mandarin the first syllable always has high pitch due to Tone 1. Additionally, in EP the contrast between the two contours is manifested by a low (L%) and rising pitch (LH%) in the second syllable; in Mandarin the second syllable bears falling pitch due to Tone 4 or a falling-rising contour due to the transition between Tone 1 and Tone 2. In other words, the melody in EP spreads over the nuclear word, whereas in Mandarin the domain of the contour is the syllable. In contrast with the findings for the native contrast (Frota et al., 2014), EP-learning infants at both 5–6 months and 8–9 months failed to discriminate this non-native contrast. These findings provide further evidence for language specific pitch perception, suggesting that intonation and lexical tone are already differentiated at 5–6 months.

In sum, work so far into early intonation suggests that there are different developmental trajectories for different types of intonation contrasts. The statement/yes-no question distinction characterized by pitch height/direction acoustic cues appears to be more salient, and thus perceived earlier, than the broad/narrow
focus contrast that is defined by a pitch timing cue. These findings highlight the importance of the nature of the cues that signal a given contrast, and may be modulated by language experience. Recent work into non-native perception of intonation suggests that infants’ perception is indeed influenced by language experience, and that language-specific perception for pitch, and for the tone/intonation distinction, emerges as early as 4–5 months of age. The findings related to the perception of intonation patterns in the presence of segmental variability show an influence of language experience earlier in development than for the perception of vowels, consonants, or even lexical tone.

Further work is required in order to better understand how the developmental trajectories of different intonation contrasts proceed. To date, only two types of intonation contrasts have been studied, and more work is needed to examine these and other intonation contrasts conveying phrasal meaning distinctions in other languages. Additionally, more work is needed to better define the aspects of infants’ particular language experience behind precocious sensitivity to pitch differences across languages.

Emerging intonation in production

Previous studies on intonational development from the viewpoint of production have followed a contour-based approach, describing the overall shapes (e.g., falling, rising, level) of the contours first produced by the child (see Snow & Balog, 2002, for a review). Studies of early intonation within the AM framework are recent. These studies address the structural properties of the melodies produced, and ask when and how young children acquire the intonational system of their ambient language. The acquisition of the intonational system involves mastering the phonology and phonetics of intonation: the former relates to the phonological inventory of pitch accents and boundary tones, and their respective meanings in language grammar and use; the latter concerns the realization of tonal patterns or their phonetic implementation, namely the alignment of pitch events with the text as well as their scaling properties. Both the phonology and the phonetics of intonation may vary in language-specific ways. In some languages, for example, alignment and/or scaling features are used phonologically as part of the system of intonation contrasts that contribute to meaning (as in the case of the focus-related pitch timing contrast in EP, described in Section 2.1).

In this section, we review the results reported in studies of the acquisition of the intonational system, and discuss similarities and differences across languages in the path of early intonational development and its relation to other aspects of grammatical development.
Acquiring the phonological inventory of tonal events

The use of the same framework of intonational analysis in various studies and for different languages allows for cross-linguistic comparison of early intonation patterns in relation to the adult intonation system. Results from several languages show that the phonological inventory of nuclear pitch accents and boundary tones is acquired early in development.

Chen and Fikkert (2007), using spontaneous speech data of three Dutch-learning children between 16 and 25 months, report that the full set of Dutch nuclear pitch accents and boundary tones is mastered by 24 months, but not before the onset of the two-word stage. Although intonational development seems to be evident relatively early in Dutch, findings for Catalan, European Portuguese and Spanish show an even earlier development of intonation. Studies on these languages, also using longitudinal corpora of spontaneous speech, all reported that children display an adult-like use of the distinct nuclear contours of their native language from the onset of speech, and appropriately use the different tunes for specific pragmatic meanings (Prieto & Vanrell, 2007, and Prieto, Estrella, Thorsson, & Vanrell, 2012, for Spanish and Catalan; Frota & Vigário, 2008 September, and Frota, Cruz, Matos, & Vigário, 2016, for European Portuguese). For example, the same segmental word can be produced as a one-word utterance expressing a request, an order, or a question, depending on the tune used (Figure 3). Early intonational development in these languages was thus found to be unrelated to the onset of combinatorial speech.

Yes-no questions have been shown to be among the sentence types that display a high degree of variation in nuclear contours, across languages and language varieties (e.g., Frota & Prieto, 2015b; Savino, 2012). In addition to cross-linguistic variation in the intonation of yes-no questions that convey a similar pragmatic meaning, yes-no questions also display a wealth of pragmatic meanings that may...
be encoded intonationally in language-specific ways (such as information-seeking, confirmation-seeking, or echo questions). In a study on the development of interrogative intonation in Catalan and Spanish-learning children between 17 and 28 months, Thorson, Borras-Comes, Crespo-Sendra, Vanrell and Prieto (2014) found that children before the two-word stage produce several types of interrogatives, and do so already reflecting the adult inventory patterns. Furthermore, there is a relationship between the type of nuclear pitch configuration produced and the pragmatic meaning of the question that agrees with the adult-like form-meaning relationships. These findings are in line with cross-linguistic evidence on the early production of the language-specific phonological inventory of nuclear pitch accents and boundary tones, and their respective meanings. They also confirm that this early development of the intonational system of the ambient language occurs regardless of the onset of combinatorial speech.

Further steps in the development of the intonational system are related with the use of prenuclear pitch accents, that is, the pitch accents that precede the nuclear contours in two-word and multiword utterances. Again, languages may differ in the types and distribution of pre-nuclear pitch accents (Frota & Prieto, 2015b; Jun, 2014b). In Dutch-learning children, the use of prenuclear pitch accents is not yet adult-like by the age of two, showing a later acquisition than nuclear pitch accents and boundary tones (Chen & Fikkert, 2007). Similar findings are reported for German and Spanish monolinguals, who show adult-like patterns by 36 months (Lleó, 2016; Lleó, Rakow, & Kehoe, 2004).

The placement of pitch accents, including the nuclear pitch accent, is a further aspect that needs to be acquired. In some languages, like English or Dutch, pitch accent placement is strongly governed by semantics and information status, and thus words may show deaccenting (or absence of pitch accent) depending on their semantic properties and on discourse context (Ladd, 2008). That is not the case in European Portuguese, for example, a language characterized by a sparse pitch accent distribution, with very few prenuclear accents, and where nuclear accent placement is final in broad focus utterances (Frota, 2014). The use of pitch accent placement by Dutch-learning two-year olds is not yet target-like, with children typically pitch-accenting both words in two-word utterances (Chen & Fikkert, 2007). In European Portuguese, a language with different pitch accent placement rules from Dutch, a tendency to pitch accent every word was also found before 24 months (Frota, Cruz, Matos, & Vigário, 2016). This tendency was interpreted as a stage in the development of prosodic phrasing, where each word forms a phrase, as also shown by other prosodic cues such as pitch reset and pause distribution. Words become integrated within the same prosodic phrase just before 24 months, and in the case of the EP-learning children studied, well before the onset
of combinatorial speech. The findings from Dutch and European Portuguese may suggest a common cross-linguistic pattern in the development of prosodic phrasing, which constrains pitch accent distribution independently of the language-specific rules that govern accent placement.

By and large, the findings available point to a precocious development of the intonational system of the ambient language (albeit with differences across languages): the inventory of nuclear pitch accents and boundary tones is adult-like close to the beginning of speech, together with the form-meaning relationships of the distinct tunes, and further developments in the use of prenuclear accents and pitch accent distribution evolve quickly and seem to be in place by the end of the second year of life. Notably, there is convergent cross-linguistic evidence that emerging intonation and its early development are independent of the onset of combinatorial speech. In Section 2, it was shown that infants are sensitive to aspects of the intonation of their ambient language from 4–5 months of age, and that they are able to linguistically use pitch cues from 6 months, for example to guide word segmentation. Moreover, infants’ perception of a given intonation contrast is a pre-requisite for the acquisition of the grammar and pragmatics it conveys. Thus, findings on infants’ perception revealed abilities of prosodic analysis of the speech signal that might help the infant to acquire the lexicon and to bootstrap aspects of the grammatical organization of the native language. It is therefore not surprising that key aspects of the intonational system develop so early in child speech. Like for perception, the findings from production are compatible with the hypothesis that prosody precedes and might promote (further) grammatical development.

Acquiring the phonetics of intonation

The early mastering of crucial aspects of the phonology of intonation does not necessarily entail that their phonetics is also acquired. Indeed, results from different languages support the view that the phonological patterns are in place before the phonetics of tonal alignment and tonal scaling become target-like. Differences across languages in the acquisition of the phonetics of intonation have also been reported, suggesting that the discrepancies between adult and early child language are not simply due to neuromotor constraints on early speech production (DePaolis et al., 2008).

Chen and Fikkert (2007) reported differences in tonal alignment of falling pitch accents (namely, H*L) between early child Dutch (from 16 to 25 months) and adult Dutch. Similarly, findings from European Portuguese show a tendency towards late peak alignment in falling nuclear accents in the beginnings of child speech, which is especially evident in the case of the H + L* pitch accent (Frota, Cruz, Matos, & Vigário, 2016). In EP, tonal alignment approached the adult-like pattern around
21 months. The fine-grained realization of alignment patterns in this language is important given the phonological use of tonal alignment in the H + L*/H* + L pitch accent contrast which underlies the distinction between broad focus and narrow focus in statements (see Section 2.1). Children have the ability to produce the difference in association and alignment between the two bitonal accents as early as 21 months, and thus well before the onset of combinatorial speech. Unlike for Dutch and EP, Spanish- and Catalan-learning children have been reported to master tune-text alignment from the beginning of speech (Prieto et al., 2012). Differently from previous studies that were based on spontaneous child speech, Astruc, Payne, Post, Vanrell, & Prieto (2013) designed a study to examine the phonetic detail of the realization of tonal targets between 24 months and six years of age in three different languages – Spanish, Catalan and English, using materials elicited by means of a controlled naming task. Their results show precise realization of tonal targets from an early age, together with differences across languages: Spanish-learning children already displayed adult-like alignment, especially of high targets, by 24 months, and earlier than Catalan- and English-learning children whose alignment patterns became increasingly target-like as they grow older. The differences across languages that have emerged in these studies strongly suggest an effect of the ambient language on the acquisition of tonal alignment patterns.

Like tonal alignment, tonal scaling was also measured in early child speech with different results across languages. The scaling of nuclear contours was initially adult-like in Catalan, Spanish, and European Portuguese children, with the low boundary tone of statements (L%) first produced as a kind of mid tone (Frota, Cruz, Matos, & Vigário, 2016; Prieto et al., 2012). Results from EP show a rapid development of scaling: by 18 months, the scaling of L% is already target-like, with other boundary tones (such as !H%) becoming target-like in their scaling between 20 and 22 months. Thus, a fine control of scaling seems to begin to evolve very early, even before alignment. The development of scaling, however, seems to succeed the mastering of alignment in Spanish and Catalan (Lleó et al., 2004; Prieto et al., 2012). Spanish-learning children do not show target-like scaling of nuclear and prenuclear accents in statements and yes-no questions before 36 months (Lleó & Rakow, 2011), whereas alignment is close to the adult-target much earlier. Nevertheless, Spanish, Catalan and English 24-month-olds have been shown to be able to control pitch scaling across words with different stress patterns, thus adjusting to the space available for the tonal movement to unfold (Astruc et al., 2013). Thus, we need to conclude that overall the phonetics of intonation is acquired early, and detailed prosodic information precedes fine-grained segmental information in acquisition, given that the phonetics are already close to the adult model when children still use truncation, simplification, or metathesis to deal with challenging segmental and word structures.
The language-specific properties that may influence the acquisition of the phonetics of intonation have yet to be examined. For example, in EP, the intonational grammar of scaling is simpler than the grammar of alignment, as scaling differences are not explored and encoded in the intonation system, unlike alignment differences (Frota, 2014). By contrast, in Spanish and Catalan the intonational grammar of scaling is more complex than in EP, with a variety of nuclear pitch accents that contrast precisely in the scaling of the peak (H + L* and ¡H + L*, L + H* and L + ¡H*), as well more complex scaling contrasts in boundary tones (e.g., LH% and L!H% – Hualde & Prieto, 2015; Prieto, 2014). This seems to suggest that phonological complexity impacts on the acquisition of tonal implementation, with more complex tonal dimensions being acquired later than simpler ones. This suggestion is in line with findings about the prosodic marking of focus in child speech in languages like Dutch and Swedish. In Dutch, prosodic focus marking involves the use of accent placement, deaccenting, choice of pitch accent type, and the particular phonetics of focal accents (Chen, 2011). In Swedish, it involves the interaction between lexical pitch accents and phrasal intonation, namely a prominence-marking high tone (Bruce, 2007). In both languages, using semi-spontaneous productions elicited by means of a picture-matching game, the phonology of prosodic focus was found to be about right by four-five years of age, with the phonetics of focus marking developing later than the phonology (Chen, 2011; Chen, this volume; Romøren & Chen, 2015). These studies of focus marking in child language provide further support for results on the development of intonation in the first years of life, showing that the phonological patterns are in place before the phonetics match the adult language.

Conclusion

This chapter focuses on early development of intonation, drawing on recent findings from perception and production, for several languages. Most of the studies reviewed have used the AM framework of intonational analysis, allowing for cross-linguistic comparison of early intonation patterns in relation to the adult intonation system.

The recent growing body of literature on infants’ perception of intonation has shown that discrimination abilities differ as a function of the types of intonation contrasts (manifested as pitch direction, pitch height, or pitch alignment distinctions), and according to the ambient language. The first finding revealed the importance of the cues that signal a given contrast, and the potential impact for language acquisition of the presence or absence of early discrimination abilities related to a specific intonation contrast that cues a given linguistic distinction (e.g., sentence types, information
structure, phrasing). The second finding highlighted that, alongside a well-documented precocious sensitivity to prosody, there is an effect of language experience in the perception of intonation patterns earlier in development than for the perception of vowels, consonants, or even lexical tone. These precocious perception abilities in the first year of life, and their attunement to the ambient language, set the stage for emerging intonation in production throughout the second year.

On the production side, cross-linguistic evidence points to an early development of the intonational system of the ambient language (albeit with differences across languages in relation to the developmental path): the phonological inventory of nuclear contours is adult-like close to the onset of speech, together with the form-meaning relationships of the distinct tunes, and further developments in the use of prenuclear accents and pitch accent distribution evolve quickly and seem to be in place by the end of the second year of life. Although the acquisition of the phonetics of tonal alignment and tonal scaling lags behind that of phonological patterns, several aspects of the phonetics of intonation are also acquired by 24 months, preceding fine-grained segmental information. Furthermore, there is convergent cross-linguistic evidence that emerging intonation, and its early development, are independent from the onset of combinatorial speech.

Findings from early perception and production of intonation suggest a link between the two, that supports precocious development. Infants’ perception revealed abilities of prosodic analysis of the speech signal that might help the infant to acquire the lexicon and to bootstrap aspects of the grammar of the native language. It is therefore not surprising that key aspects of the intonational system develop early in child speech. As for perception, the findings from production are compatible with the hypothesis that prosody precedes and might promote (further) grammatical development. There are aspects of intonational development, however, that are acquired later than the first years of life, as in the case of focus development (Section 3.2; Chen, this volume; Ito, this volume), or the relationship between speaker belief states and prosody (Armstrong & Hübscher, this volume). Even if intonational development naturally proceeds with age, the perception and production abilities acquired early are crucial to foster further language development (see, also, Bhatara, Boll-Avetisyan, Höhle, & Nazzi, this volume; de Carvalho, Dautriche, Millotte, & Christophe, this volume).

The overall finding of differences across languages in the developmental trajectories of early intonational acquisition suggests that perception is not simply guided by (domain-general) abilities, and early production is not only driven by neuromotor constraints, which would entail similar paths. However, more work is needed to better define the language-specific properties that may influence infants’ sensitivity to pitch differences across languages, as well as differences in the acquisition of the phonology and phonetics of intonation revealed in production.
Despite recent advances, the current understanding of how intonation is acquired is still fragmentary. The AM framework of intonation analysis has offered an useful tool to investigate the development of intonation in perception and production, but only a few languages have been examined so far. The fact that, in the languages covered, only a handful of intonation contrasts, and their realizations, have been investigated is a further limitation. It is also important to note that all studies to date that have examined infants’ perceptual abilities are discrimination studies. The lack of identification studies in this area, although probably motivated by the young ages observed, is a particularly relevant issue that needs to be acknowledged and addressed in future research. Clearly, more data and, in particular, more controlled experiments are needed. These could take advantage of ‘old’ behavioral methods in combination with newer techniques, such as event-related potentials (ERPs), near infrared spectroscopy (NIRS), or eye-tracking, all of which may be used to assess early stages in development and hold the promise to offer new insights in this exciting field of research.

Understanding the developmental path of intonation towards a full adult-like competence will not only impact on our knowledge of intonation itself, but contribute to our understanding of the acquisition of other linguistic categories and structures, given the potentially bootstrapping role of intonational cues (Höhle, 2009). Furthermore, as a lack of sensitivity to intonation, together with its atypical production, may indicate low social skills and developmental disorders (Koegel et al., 2010; Peppe, this volume), it will also contribute towards tracing the development of communication and social interaction, and signal language impairments.

**Funding**

This research was supported by grants PTDC/CLE-LIN/108722/2008, EXCL/MHCLIN/0688/2012, PTDC/MHCLIN/3901/2014, and IF/01614/2012 from the Foundation for Science and Technology (Portugal) and Fundo Social Europeu – Programa Operacional Potencial Humano.

**Acknowledgements**

The authors would like to thank the reviewers and the editors of this volume for comments on previous versions of this chapter. The authors also gratefully thank Marina Vigário, Marisa Cruz, Cátia Severino, Susana Correia, Nuno Matos and Simão Cortês for help at different stages of the present work. Special thanks are due to the children, families and nurseries that have collaborated in our studies.
References


Author Queries

Please provide a complete reference for the citation ‘(Frota, Butler, Lu, and Vigário (2016))’ of this chapter.

Please provide a complete reference for the citation ‘(Nazzi, Bertoncini, & Mehler, 1998)’ of this chapter.

Please provide a complete reference for the citation ‘(Frota, Cruz, Matos, & Vigário, 2016)’ of this chapter.