

To appear in *Infancy*

Infancy, 1–19, 2013

Copyright © International Society on Infant Studies (ISIS)

ISSN: 1525-0008 print / 1532-7078 online

DOI: 10.1111/inf.12037

This is the manuscript version.

Infants' perception of intonation: Is it a statement or a question?

Sonia Frota

Joseph Butler

Marina Vigário

University of Lisbon

Corresponding author:

Joseph Butler

Lisbon Baby Lab (CLUL)

Dept. Linguística, Faculdade de Letras,

Universidade de Lisboa,

Alameda de Universidade,

1600-214 Lisboa

Portugal

jbutler@fl.ul.pt

+351 217920052

Key words: Statement intonation; question intonation; language acquisition; infant perception; prosodic discrimination

Abstract

The ability to distinguish phonetic variations in speech that are relevant to meaning is essential for infants' language development. Previous studies into the acquisition of prosodic categories have focused on lexical stress, lexical pitch accent, or lexical tone. However, very little is known about the developmental course of infants' perception of linguistic intonation. In this study, we investigate infants' perception of the correlates of the statement/yes-no question contrast in a language that marks this sentence type distinction only by prosodic means, European Portuguese (EP). Using a modified version of the visual habituation paradigm, EP learning infants at 5-6 and 8-9 months were able to successfully discriminate segmentally varied, single-prosodic word intonational phrases presented with statement or yes-no question intonation, demonstrating that they are sensitive to the prosodic cues marking this distinction as early as 5 months, and maintain this sensitivity throughout the first year. These results suggest the presence of precocious discrimination abilities for intonation across segmental variation, similarly to previous reports for lexical pitch accent, but unlike previous findings for word stress.

(word count = 173)

Infants' perception of intonation: Is it a statement or a question?

Introduction

Infants' ability to distinguish between forms of phonetic variation in speech that are relevant to meaning is essential for their language development. Languages differ in how the distinction between statements and yes-no questions is conveyed. For example, in European Portuguese, this distinction is marked by prosodic means (Frota, 2002), and so infants must distinguish between the prosodic features associated with statements and yes-no questions to be able to understand the implied meaning of the utterance. In the present study, we aimed to investigate how the development of infants' perception of linguistic intonation unfolds, using the statement/yes-no question prosodic contrast in European Portuguese. Following Gussenhoven (2007) and Ladd (2008), intonation is understood as the use of prosodic features to express sentence-level meanings. Sentence type distinctions, such as the statement/yes-no question contrast, are among the phrasal meanings that may be conveyed by intonation.

Infants' early sensitivity to the prosodic properties of speech is well documented. Studies on newborns' prosodic abilities have shown that they are sensitive to the global rhythmic properties that differentiate between languages (Nazzi, Bertoncini & Mehler, 1998), to word-level rhythm, that is to stress patterns (Sansavini, Bertoncini & Giovanelli, 1997), and to differences in the pitch contours of words resulting from contrasting lexical pitch accents (Nazzi, Floccia & Bertoncini, 1998). These studies suggest that newborns' sensitivity to prosody may not depend on exposure to a specific language. However, they leave open the possibility that this very early sensitivity may evolve differently according to the specific prosodic properties of the native language.

Studies into the acquisition of sound categories have shown a developmental change before the infant's first birthday, whereby discrimination of non-native contrasts declines giving way to discrimination abilities specifically attuned to the native language (Werker & Tees, 1984; Kuhl, 2004; Saffran, Werker & Werner, 2006). It is also known that this general developmental pattern may be

modulated by perceptual assimilation differences and by the phonetic salience of non-native and native contrasts (Best & McRoberts, 2003; Narayan, Werker & Beddor, 2010). For native contrasts, at least two different paths seem to be possible: an initial sensitivity in the first half of the first year of life that is maintained in the second half (Polka, Colantonio & Sundara, 2001), or an initially poor sensitivity followed by sensitivity gains (Kuhl, Stevens, Hayashi, Deguchi, Kiritani & Iverson, 2006; Narayan et al., 2010). For non-native contrasts, a decline in discrimination abilities is usually reported (Werker & Tees, 1984; Best & McRoberts, 2003; Kuhl, 2004). However, most previous studies have focused on examining segmental categories (i.e., vowels and consonants), and not prosodic categories. Thus much less is known about how the perception of prosodic contrasts develops in the first year of life. The few studies on the development of infants' perception of prosodic categories have almost exclusively investigated the prosodic properties that can be used to distinguish lexical items, namely lexical stress, lexical tone and lexical pitch accent.

Regarding lexical stress, infants are able to discriminate word stress patterns if segmental variation in the stimuli is limited (a single token with stress-initial and stress-final realizations, or different tokens in which consonants, but not vowels, are varied), irrespective of whether stress is used contrastively or not in the input language. This has been shown for Italian newborns (Sansavini, Bertoncini & Giovanelli, 1997), for German-learning and French-learning infants at 4 and 6 months (Friederici, Friedrich & Christophe, 2007; Höhle, Bijeljac-Babic, Herold, Weissenborn & Nazzi, 2009; Skoruppa, Pons, Bosch, Christophe, Cabrol & Peperkamp, 2013), Spanish-learning infants at 6 months (Skoruppa et al., 2013), and French-learning infants at 9 and 10 months (Skoruppa, Pons, Christophe, Bosch, Dupoux, Sebastián-Gallés, Limissuri & Peperkamp, 2009; Bijeljac-Babic, Serres, Höhle & Nazzi, 2012). Interestingly, bilingual infants learning French, a language where stress is not contrastive, and another language which uses stress contrastively (like Spanish or English), show enhanced sensitivity to stress contrasts relative to French monolinguals, in line with the general patterns usually found for native *versus* non-native contrasts (Bijeljac-Babic et al., 2012). A different picture emerges when segmentally varied stimuli, closer to the phonetic

variability found in speech, are used (e.g., lists of stimuli that vary in both consonants and vowels). In these conditions, Italian newborns do not discriminate between initial and final stress (Sansavini, 1997), French and Spanish-learning 6-month-olds do not discriminate stress patterns (Skoruppa et al., 2013), and 6-month-old English-learning infants show no preference for the predominant stress pattern of the native language (Jusczyk, Cutler & Redanz, 1993). However, by 9 months both Spanish and English infants are able to discriminate word stress patterns, and the latter already show a preference for the predominant stress pattern of English (Skoruppa et al., 2009; Skoruppa, Cristià, Peperkamp & Seidl, 2011; Jusczyk et al., 1993). Only French infants, who are learning a language where stress is not contrastive, show no successful discrimination at 9 months (Skoruppa et al., 2009). These findings suggest that word stress discrimination abilities in the presence of segmental variation are learned by attuning to the native language prosodic system, and thus the developmental path for word stress perception involves sensitivity gains during the first year of life.

Another prosodic property that can be used to distinguish word meanings is lexical tone. Lexical tones are found in languages like Mandarin, Cantonese or Thai (which are thus tone languages), but not in English or French. Studies on infants' perception of lexical tone have observed that English and French-learning infants show a decline in tone discrimination from 4 to 9 months of age, whereas learners of tone languages show stable discrimination abilities in the first year of life, from as early as 4 months (Mattock & Burnham, 2006; Mattock, Molnar, Polka & Burnham, 2008; Yeung, Chen & Werker, 2013). In these studies, tone perception was examined in the absence of variability (a single syllable-size token produced with the contrasting tones). The developmental path found for tone discrimination is thus in contrast with stress discrimination abilities under absence of variability: in the case of stress, discrimination is maintained regardless of the role of stress in the native language (as reported above); in the case of tone, discrimination abilities are only maintained by learners of tone languages. However, infants' ability to perceive tone contrasts from variable tokens seems to show learning effects similar to those found for stress perception in the presence of segmental variation: Mandarin-learning infants were able to perceive native tone contrasts in

variable contexts (variable tone in the preceding syllable) at 8 to 11 months, but not at 4 to 6 months (Shi, 2010).

Lexical pitch accents are a further type of prosodic property from the lexical prosody domain. In languages like Japanese, or Basque, pitch changes at the word-level are used to distinguish lexical meaning. Japanese-learning infants are able to discriminate the native pitch accent contrast from as early as 4 months, and their discrimination abilities are maintained during the first year of life (Sato, Sogabe & Mazuka, 2009). Importantly, for lexical pitch accent successful discrimination was obtained with variable stimuli (word lists varying in their consonants and vowels) as early as 4 months of age, a result unlike that obtained for lexical stress or lexical tone.

To sum, studies on infants' perception of lexical prosodic contrasts reveal differences in discrimination abilities for stress, tone and pitch accent. These studies also show that variability in the stimuli may affect early perception of some prosodic contrasts, but not all: In the first half of the first year discrimination of lexical stress and lexical tone seem to be negatively affected by variability, whereas discrimination of lexical pitch accent is not affected. If discrimination is successful with phonetically varied stimuli, this indicates that infants are not only able to perceive the acoustic cues for the prosodic contrasts, but are also able to extract and generalize the contrastive patterns across the phonetic variability. Given the findings for lexical prosody, we ask whether early perception of linguistic intonation, and of sentence type intonation contrasts in particular, follows a path similar to lexical stress (and perhaps tone) – protracted development of early discrimination of intonation in the presence of variability, or to lexical pitch accent – precocious discrimination abilities across phonetic variability.

To date, very few papers have begun to investigate the developmental course of infants' perception of linguistic intonation. It has long been known that infants prefer rising contours, high pitch and expanded melodies right from birth, and that they are sensitive to changes in melodic contours (e.g. Trehub, Bull & Thorpe, 1984; Papoušek, Bornstein, Nuzzo, Papoušek & Symmes, 1990), but grammatically relevant intonation contrasts have only recently begun to be examined.

Intonation, like lexical stress, lexical tone and lexical pitch accent, is a prosodic dimension that varies across languages and impacts upon meaning. However, unlike lexical stress, tone and pitch accent, which are meaningful at the word level, intonation conveys phrasal meanings, such as sentence type distinctions. These distinctions are important for language acquisition. Statements and questions are basic and frequent sentence types (Newport, Gleitman & Gleitman, 1977), and thus the ability to distinguish between them is crucial to process most sentences in the input a child is exposed to. Additionally, questions and statements are crucial for communication and social interaction (Fernald, 1989; Koegel, Koegal, Green-Hopkins & Barnes, 2010).

Across languages, the sentence type distinction between statements and yes-no questions may be marked both by morphosyntax and intonation (e.g., English, Catalan – Ladd, 2008; Prieto & Rigau, 2007), not marked by intonation (e.g., Vata, Shekgalagari, Cantonese – Rialland, 2007; Hyman & Monaka, 2011; Wong, Chan & Beckman, 2005), or cued by intonation only (e.g., Italian, Portuguese – Maiden & Robustelli, 2000; Mateus, Brito, Duarte, Faria, Frota, Matos, Oliveira, Vigário & Villalva, 2003). When intonation contributes to the distinction, the intonational features used may also vary across languages. The most common features are utterance final pitch movement (that is boundary tones), peak alignment, or changes in pitch register (Ladd, 2008). In Dryer's (2011) survey on how languages indicate interrogativity, the use of intonation only is estimated to be the second most frequent means to signal a yes-no question as different from a statement.

To our knowledge, only three studies have started to examine the discrimination of sentence type intonation contrasts by infants, and all of them in the English language. Best, Levitt & McRoberts (1991) looked at English-learning infants' discrimination of exclamations and wh-questions, both in native and non-native (Spanish) sentences. The stimuli they used differed across the two sentence types in word order (presence of inversion in wh-questions, with the auxiliary verb placed before the subject) and lexical properties (presence of a wh-word in questions), besides prosodic properties. Their results showed that 6-8 month olds discriminated the prosodic contrast both in the native and non-native conditions, but 10-12 month olds failed to discriminate it in either

of the conditions. The unsuccessful discrimination of the native contrast by older infants was paradoxical under the view that discrimination abilities evolve by attunement to the native language. On the other hand, it is not clear which cues younger infants may have used in the discrimination of the sentence type contrast, since prosodic cues (which were quite variable within each category, at least for English sentences) as well as lexical and word order cues (possibly reflected in distributional cues to which infants have been shown to be sensitive from 4 months of age – Bosch & Sebastián-Gallés, 2001; Friederici, Mueller & Oberecker, 2011) were present in the stimuli. In Geffen & Mintz (2011) study, the discrimination of the statement/yes-no question contrast was addressed using stimuli that contained both word order (inversion, as in ‘Can I play...?’) and intonation cues. The authors report that 7-month-olds successfully discriminate the two sentence types. However, as in Best et al. (1991), it is not clear whether infants used the differing prosodic characteristics of the statement/yes-no question contrast or whether they were able to utilise the available word order. Soderstrom, Ko and Nevzorova (2011) also examined English-learning infants’ perception of statements and yes-no questions. Unlike in the two previous studies, they used materials that neutralised the word order cue by having uninverted yes-no questions as their stimuli. They found infants displayed an overall preference for questions. However, their results do not provide direct evidence that infants are able to discriminate between the two categories, but only evidence for an attentional bias towards questions that could be related to a general preference for high/rising pitch repeatedly found in the literature (Trehub et al., 1984; Papoušek et al., 1990). In addition, they have used quite a broad age range of participants (4.5 to 24 months), and so it is difficult to interpret any developmental trends from their data. These few reports leave the fundamental question unanswered of how the perception of linguistic intonation, and of sentence type intonation contrasts in particular, unfolds during the first year of life.

In this study, we investigated infants’ perception of the statement/yes-no question contrast in a language that marks this sentence type contrast only by prosodic means. In the standard variety of European Portuguese (EP), the two sentence types are differentiated by the nuclear contour, that

is by the melody on the nuclear syllable and subsequent post-tonic syllable(s). This melody comprises the pitch movement realized on the stressed syllable of the nuclear word (which has falling pitch in both declaratives and questions – H+L*) and the pitch found on the final syllable (which is low in statements – L%, and rising in questions – LH%). Thus, the statement/yes-no question contrast in EP is more specifically manifested by the melodic shape at the right-edge of the utterance (statements: H+L* L%; questions: H+L* LH%). Questions also show longer duration of nuclear and post-nuclear syllables than statements (Frota, 2002). This prosodic contrast has been shown to be perceived by adult native speakers of EP (Falé & Faria, 2005).

In languages like English or EP, a given melody may be realized either on a monosyllabic one-word utterance or on longer multiword utterances without changing its key phonological features (Gussenhoven, 2004; Ladd, 2008). Therefore, the exact same melody (nuclear contour) may be applied to strings with varying numbers of syllables, as illustrated by Ladd's (2008) example of 'Sue!?' or 'A driving instructor!?' each bearing the same rise-fall-rise melody (uttered in reply to 'I hear Sue's taking a course to become a driving instructor' – Ladd, 2008: 45-46). Likewise, the EP declarative and question contours described above may appear in either single-word or longer utterances (Frota, 2000, 2002). An infant exposed to a language like English or EP is thus subjected to an input that contains melodies applied to both one-word and multiword utterances.

However, in EP as in many other (if not all) languages that use melodies to express phrasal meanings, the melody is a property of the intonational phrase, and not of utterances or sentences, although intonational phrases may (but need not) correspond to utterances or sentences (Frota, 2000; Nespor & Vogel, 2007). The intonational phrase is 'the smallest prosodic unit that can have a melody' (Ladd, 2008: 298). Although it has long been known that infant-directed speech (IDS) tends to consist of short utterances (e.g., Snow, 1972), and previous findings on English and Dutch reported that 9% to 12% of IDS utterances consist of isolated words (Brent & Siskind, 2011; Fernald & Hurtado 2006), to our knowledge prior studies on IDS have not investigated the properties of intonational phrases. In an ongoing study, we computed intonational phrases from transcripts of

spontaneous speech directed to four EP-learning infants aged between 11 and 16 months. We found that, overall, 43% of intonational phrases consisted of one prosodic word (that is, one word with a stressed syllable, which roughly corresponds to a lexical word – Vigário, 2003), and that the speech each infant had been exposed to contained between 37% to 47% of single prosodic word intonational phrases (Vigário, Frota, Martins & Butler, in progress). Furthermore, nearly half of these intonational phrases instantiated either the declarative or question sentence types (27.3% and 19.4%, respectively). We concluded that single prosodic word intonational phrases are a robust feature of infant directed speech in EP, and thus the most common prosodic domain for a melody in IDS is an intonational phrase containing one prosodic word.

In the present study, we ask how the development of infants' perception of linguistic intonation unfolds. We focus on the statement/yes-no question contrast, which is conveyed only by prosodic means in European Portuguese, and use segmentally varied stimuli where the two contrastive melodies appear in the most common prosodic domain for a melody in infant-directed speech. By examining 5-6 and 8-9 month-olds perception of this prosodic distinction in a visual habituation procedure, we address the following questions: Is the developmental path of how infants acquire the ability to discriminate sentence type intonation contrasts in their native language characterized by an initial sensitivity that is maintained, or by an initially poor sensitivity followed by sensitivity gains/learning effects?; Does phonetic variability affect early perception of intonation?; How does the developmental trajectory of the perception of native intonation contrasts compare to that of lexical prosodic categories (stress, tone and pitch accent) and to the few previous reports on intonation discrimination by English-learning infants? If discrimination abilities are maintained throughout the first year, the developmental trajectory of intonation would be most similar to that reported for lexical pitch accent. If they are a newly learned ability, only the older infants would be able to distinguish between statement and yes-no question prosody and the developmental trajectory of intonation would be most similar to that reported for lexical stress. If none of these patterns obtains, the development of infants' perception of linguistic intonation would be unlike

previous descriptions for other prosodic categories. However, if discrimination is indeed successful, and given the correspondence between contrastive prosodic features and the statement/question distinction in EP, prosody may facilitate the acquisition of the grammar of the two sentence type categories very early on in development.

Method

Participants

Forty infants participated in this study, split into two age groups; 20 younger (8 female, mean age 5 months 29 days, range 5 months 3 days – 6 months 23 days) and 20 older (10 females, mean age 8 months 12 days, range 7 months 11 days – 9 months 29 days). All were normally developing infants raised in monolingual European Portuguese homes, recruited from the wide Lisbon area. In addition, 8 infants (3 younger, 5 older) were rejected from the study due to fussiness (5), a parent who was not European Portuguese (1), an older, autistic sibling (1) and software error (1).

Materials

A set of 16 pseudo-words was used to create statement and question segmentally varied single-prosodic word intonational phrases. The pseudo-words were bisyllabic all sonorant sequences with initial stress, produced by a female native speaker in infant-directed speech (malo, lemo, loma, mela, rono, rano, nurra, nirra, lamo, milo, mola, luma, norro, reno, nerra, rina).

Mean duration and mean fundamental frequency (F0) for the pitch fall in the first syllable (H+L*), the F0 range in the second syllable (a negative value indicating that pitch falls through the syllable and a positive value indicating a rise in pitch) and the final F0 value in the contour (L% or LH%) were computed for the declarative and interrogative versions of the 16 pseudo-word utterances (see Table 1). As expected, the acoustic measurements showed that the two categories were distinguished by the boundary tone difference (L% for declaratives, final F0 range=-24.75 Hz, final F0 value=163 Hz; LH% for interrogatives, final F0 range=191.56 Hz, final F0 value =380 Hz), and

the longer duration of yes-no questions (236 ms longer on average). However, they also showed that the declarative versions tended to have a lower initial peak than the interrogative versions, and thus a smaller range for H+L* (38Hz lower on average). To ensure that the intonation contrast to be tested was carried out only on the final boundary tone difference (L% vs. LH%), the declarative versions were manipulated to always show a pitch fall identical to the questions (Praat was used to manipulate the peak value in H+L*).

Insert Table 1 here

The naturalness of this manipulation was assessed by means of a rating task where the manipulated declarative and natural interrogative versions of the stimuli (eight items in the declarative and a different set of eight items in the interrogative) were presented to 10 adult subjects, who were asked to rate each item on a scale of 1 (definitely a statement) to 5 (definitely a question). Response reaction times were also recorded. A one-sample Chi-Square Test revealed that responses were significantly different to chance for declarative ($p < .001$) and interrogative ($p < .01$) stimuli. A non-parametric independent samples Mann-Whitney test revealed significant differences between declarative and interrogative stimuli for both score ($z = 8.42$, $p < .001$, Average declarative rating = 1,4; Average question rating = 3,6) and reaction time ($z = 3.44$, $p < .01$, Average RT for declaratives = 2090; Average RT for questions = 2572). Interestingly, manipulated declaratives were more consistently rated as 'definitely a statement' than natural interrogatives were rated as 'definitely a question', and reaction times for declaratives were significantly shorter. On the basis of these results we concluded that manipulated declaratives were naturally perceived as statements, and thus the manipulated declarative versions were used as stimuli together with the interrogative versions of the 16 pseudo-words (examples of each are provided in Figure 1; sound files of the stimuli are available at

http://ww3.fl.ul.pt//laboratoriofonetica/babylab/Infants_Perception/Infants_perception_intonation_supporting_materials.htm).

Insert Figure 1 here

In addition to differences in intonation, different pseudo-words were used for habituation and test phases (half of the stimuli for each: Habituation trials – malo, lemo, loma, mela, rono, rano, nurra, nirra; Test trials – lamo, milo, mola, luma, norro, reno, nerra, rina).

Procedure

A modified version of the visual habituation paradigm (Stager & Werker, 1997) was used, a methodology that has been applied in testing discrimination in infants (Mazuka, Cao, Dupoux & Christophe, 2011; Sato et al., 2009). Infants were seated on the caregivers lap, facing a computer monitor displaying a colourful, attractive image. Each trial began once the infant fixated this image. The infant was then presented with a red and black checkerboard display, together with the habituation stimuli delivered through speakers hidden behind the monitor. The stimuli for each trial was a sound file consisting of 8 intonational phrases, and the length of each sound file was 16 seconds (due to the differing durations of declarative and yes-no question stimuli, the inter-stimulus-interval was adjusted to ensure all sound files were 16 seconds total in length, with an ISI of 1471 ms on average for declarative trials and 1235 ms for interrogative trials). The trial continued until the sound file completed, and the image then reverted to the colourful image. When the infant fixated the image, the next trial started. Presentation of stimulus was controlled by the LOOK software (Meints & Woodford, 2008), and an experimenter who observed the infant through a camera discretely placed above the monitor, and who began a trial by pushing a button when the infant fixated the image, and kept the button pressed while the infant oriented towards the

checkerboard image. Both the experimenter and caregiver wore music-playing headphones to mask the experiment sounds, and the experimenter was unaware of the experimental conditions. The LOOK software recorded and monitored total looking times to each trial, and habituation trials continued until a pre-set habituation criteria was reached, i.e. average looking time to the last four habituation trials heard was less than 60% of the average looking time to the first four habituation trials. Infants were then presented with two test trials, one “same” (as the habituation) and one “switch” (different to the habituation). Half of the infants habituated with statement intonation, half with question intonation, and the order of presentation of the test trials was counterbalanced (same/switch first). If infants were sensitive to the prosodic contrast, they should display longer looking times to the switch trials.

Results

The habituation phase was analysed first. Overall, average looking times to the first 4 trials (mean = 12.46, SD = 2.7) were longer than the average looking times to the last 4 trials (mean = 7.46, SD = 2.57), and this pattern was also true for younger (first 4 trials mean = 13.27, SD = 2; last 4 trials mean = 8.06, SD = 2.44) and older (first 4 trials mean = 11.65, SD = 3.05; last 4 trials mean = 6.86, SD = 2.58). The average looking times to the first 4 and last 4 habituation trials were examined using a repeated measures Analysis of Variance (ANOVA) with a within-participant factor of habituation (beginning of habituation phase vs. end of habituation phase) and two between-participant factors of age group (younger vs. older) and habituation condition (declarative vs. interrogative). There was a significant effect of habituation ($F(1,36) = 629.71, p < .001, \eta^2 = .95$) and a significant effect of age group ($F(1,36) = 10.98, p < .01, \eta^2 = .23$), but no significant effect of type of habituation condition ($F(1,36) < 1$), and no significant interactions. An independent sample t-test comparison across the younger and older groups of infants revealed no significant difference in the number of trials needed to habituate (younger mean = 12.25, older mean = 10.9, $t(38) = 1.02, p = .31$), however younger infants displayed significantly longer looking times to the first four habituation trials than older infants (younger mean = 13.27, older mean = 11.65, $t(38) = 2.59, p < .05$). The significant differences

between the age groups suggest that there are differing levels of attention, with older infants displaying overall less attention in this task. However, the lack of significant interactions suggests that both groups performed similarly in habituating to the stimuli, again with no difference between statement and yes/no question intonation.

Insert Figure 2 about here

Average looking times for same and switch trials in the test phase by age group are presented in Figure 2. As can be seen, both groups display longer looking times to the “switch” test trials (younger same mean – 6.88, SD – 2.18, switch mean – 10.86, SD – 3.13; older same mean – 6.53, SD – 3.2, switch mean – 8.93, SD – 2.67). A repeated measures ANOVA, with a within-participant factor of trial type (same vs. switch) and two between-participant variables of age group (younger vs. older) and habituation condition (declarative vs. interrogative), revealed a significant effect of trial type only ($F(1,36) = 54.18, p < .001, \eta^2 = .6$). Neither the main effect of age group ($F(1,36) = 2.13, p = .15, \eta^2 = .06$), nor of habituation ($F(1,36) = 2.02, p = .16, \eta^2 = .05$) were significant, nor were any of the interactions (trial type x age group, $F(1,36) = 3.29, p = .08, \eta^2 = .08$; all the other interactions, $F(1,36) < 1$). Paired t-tests for each age group separately revealed significant differences between same and switch trials for both age groups (younger $t(19) = 6.1, p < .001, d = 1.474$; older $t(19) = 4.42, p < .001, d = 0.816$). Overall, 18 out of 20 younger and 17 out of 20 older infants had longer looking times to the switch trials.

Discussion

In this study, we examined 5-6 and 8-9 month-old European Portuguese learning infants’ perception of the statement/yes-no question prosodic distinction to investigate how the development of infants’ perception of linguistic intonation unfolds in the first year of life. Our results demonstrate that infants at both ages are able to successfully discriminate utterances that differed

only in the prosodic features associated with statements and questions. These findings suggest that the developmental path of how infants acquire the ability to discriminate sentence type intonation contrasts in their native language is characterized by an initial sensitivity that is maintained, and not by an initially poor sensitivity followed by sensitivity gains/learning effects. Given that infants were presented with phonetically varied stimuli, discrimination is assumed to reflect infants' ability to extract the common prosodic features that characterize each of the sentence types. This discrimination ability is a pre-requisite for the acquisition of the statement/yes-no question distinction in a language that marks this sentence type contrast only by prosodic means. The ability to perceive the differing intonational properties of statements and questions very early on in development may thus facilitate the acquisition of the two sentence type categories.

Considering earlier work on infants' perception of lexical prosodic contrasts, our results approximate the developmental trajectory of intonation perception to that of lexical pitch accent and tone (Sato et al., 2009; Mattock & Burnham, 2006; Yeung et al., 2013), in contrast to stress (Höhle et al., 2009; Bijeljac-Babic et al., 2012; Skoruppa et al., 2013): for intonation, lexical pitch accent and tone, stable discrimination abilities were found very early on in development for learners of languages that use the studied prosodic contrasts in their grammars. In addition, early discrimination abilities for both intonation and lexical pitch accent are not hindered by variability in the stimuli, unlike in the case of stress (the effects of high phonetic variability have not yet been studied for lexical tone, but see Shi, 2011). Taken together, these findings uncover differences in discrimination abilities across prosodic contrasts that seem to group pitch-based (or pitch-dominant) contrasts as those that tend to show precocious discrimination abilities. Although much more research is needed – in particular studies exploring the effect of segmental variability in lexical tone perception and studies of other intonation contrasts –, we would like to suggest that phonetic salience and distribution across languages may link to these differences. It is known that phonetic salience plays a role in the development of perceptual sensitivity (Narayan et al., 2010; Yeung et al., 2013), and pitch contrasts seem to be especially salient to infants and are detected earlier than

other sound features (Bion, Benavides-Varela & Nespor, 2011; He, Hotson & Trainor, 2007). In addition to phonetic salience, some kind of linguistic melody seems to be present in all languages (whether in the form of tone, lexical pitch accent, or intonation – Gussenhoven, 2004). However, the same is not true for stress, since there are languages with no apparent phonetic stress (Beckman, 1986). Pitch contrasts are thus highly salient acoustic distinctions that are very common across languages, unlike stress contrasts.

To our knowledge, only one study has previously examined infants' perception of linguistic intonation (the statement/yes-no question contrast in English) using stimuli that varied only in their intonational properties, that found no evidence for discrimination and a general attentional bias for questions (Soderstrom, Ko & Nevzorova, 2011). Several reasons may explain the differences between our results and those reported by Soderstrom et al. (2011). First, statements and yes-no questions in English are distinguished both by morphosyntax (inversion, that is the presence of an auxiliary or operator before the subject in questions) and intonation (rising intonation in questions), whereas in European Portuguese the sentence type contrast is conveyed by intonation only: for example, the statement 'He likes Mary' (in EP, 'Ele gosta da Maria'), has the yes-no question counterpart 'Does he like Mary?' (in EP, 'Ele gosta da Maria?', but not 'Gosta ele da Maria?'). Although uninverted yes-no questions are also possible in English, it may well be the case that there are different developmental trajectories between infants learning languages with differing cues to this sentence type contrast. Second, studies on the possible meanings of intonational forms suggest that the correspondence between intonational form and sentence type distinctions is not as robust in English as in EP (Bolinger, 1989; Pierrehumbert & Hirschberg, 1990; Frota, 2002, in press). In fact, Soderstrom et al. report variable contour shapes in their stimuli (especially for the declaratives). Other reasons may be methodological. The participants in the Soderstrom et al. study have quite a broad age range (4.5 to 24 months) and this may have masked any developmental trends. Moreover, the stimuli used were multiword sentences, and their complexity in terms of prosodic structure (namely, the number of intonational phrases or melody domains) is not described. Future work is

needed to tease apart these factors, by testing younger and older English-learning infants' perception of this sentence type contrast using less complex stimuli, as well as the perception of infants learning other languages that use different cues for the contrast (or that weight the cues differently). Furthermore, we have only tested EP-learning infants' perception of the sentence type contrast in single-prosodic word intonational phrases and thus the question of whether and when EP infants are able to discriminate intonation patterns in multiprosodic word intonational phrases remains to be explored. Based on previous work showing that intonational phrase boundaries are especially salient to infants and that they attend differently to words aligned with such boundaries, i.e. to the words that carry the nuclear contour, as early as 6 months (Shukla, White & Aslin, 2011), we would expect to replicate the presently observed pattern.

Another question that should be examined in future work is whether successful discrimination of native sentence type intonation contrasts emerges at 5 months, even earlier in development, or is present from birth. Furthermore, we must ask whether the discrimination abilities across phonetic variability that we have found for both 5-6 and 8-10 months-olds can be linguistically interpreted in similar ways. Failure to discriminate in the presence of variability (together with successful discrimination in the absence of variability) has been generally attributed to the ability to perceive the acoustic cues for the contrasts, but not to extract and generalize across phonetic variability the property shared by each type of variable stimuli. The former ability is usually related to the detection of low-level acoustic differences, whereas the latter has been interpreted as suggesting processing at the phonological level (Mazuka et al., 2011; Skoruppa et al., 2013). However, Sato et al., (2009), studying the perception of lexical pitch accent by Japanese infants, demonstrated that infants were able to discriminate the lexical pitch accent contrast both at 4 and 10 months (in a visual habituation procedure), but brain responses to the contrast differed between 4 and 10 months-olds (in a NIRS paradigm). Only the older infants show a different brain response pattern for a change in lexical pitch and a change in pure tones (i.e., non-linguistic stimuli). The authors thus argue that only older infants are processing the native contrast as linguistic.

In conclusion, the present findings that European Portuguese learning infants are able to discriminate the intonation contrast marking declaratives and yes-no questions as early as 5 months in the presence of phonetic variability, and maintain this discrimination ability throughout the first year, add to our knowledge of how the perception of prosodic contrasts develops by extending previous research to the investigation of intonation contrasts. As the present study is the first to show successful discrimination of an intonation contrast on the basis of prosodic cues only, it is not known whether the current findings extend to other languages where the statement/yes-no question contrast is overtly cued solely by intonation. It is a task for future research to examine this and other intonation contrasts conveying phrasal meaning distinctions, in Portuguese and other languages.

Acknowledgments

This work was supported by Grant PTDC/CLE-LIN/108722/2008 from the Foundation for Science and Technology (Portugal). The authors would like to thank Cátia Severino for help with data collection. Special thanks are due to the families and nurseries that have collaborated in this study.

References

- Beckman, M. E. (1986). *Stress and non-stress accent*. Dordrecht: Foris.
- Best, C. T., Levitt, A. & McRoberts, G. W. (1991). Examination of language-specific influences in infants' discrimination of prosodic categories. *Proceedings of the XIIth International Congress of Phonetic Sciences, Vol. 4* (pp. 162-165) Aix-en Provence: Universite de Provence, France.
- Best, C. C. & McRoberts, G. W. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language & Speech, 46*, 183-216.
- Bijeljac-Babic, R., Serres, J., Höhle, B. & Nazzi, T. (2012). Effect of bilingualism on lexical stress

- pattern discrimination in French-learning infants. *PLoS ONE*, 7(2): e30843.
doi:10.1371/journal.pone.0030843.
- Bion, R. A. H., Benavides-Varela, S. & Nespor, M. (2011). Acoustic markers of prominence influence infants' and adults' segmentation of speech sequences. *Language and Speech*, 54(1), 123-140.
- Bolinger, D. (1989). Intonation and its uses. Stanford University Press.
- Bosch, L. & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, 2(1), 29-49.
- Brent, M. R. & Siskind, J. M. (2011). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, 2(1), 29-49.
- Dryer, M. S. (2011). Polar Questions. In: Dryer, Matthew S. & Haspelmath, Martin (eds.)
The World Atlas of Language Structures Online. Munich: Max Planck Digital Library, chapter 116. Available online at <http://wals.info/chapter/116> . Accessed on 2012-10-09.
- Falé, I. & Hub Faria, I. (2005). Intonational contrasts in EP: a categorical perception approach. *INTERSPEECH-2005*, 1705-1708.
- Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development*, 8, 1497-1510.
- Fernald, A. & Hurtado, N. (2006). Names in frames: infants interpret words in sentence frames faster than words in isolation. *Developmental Science*, 9(3), F33-F40.
- Friederici, A. D., Friedrich, M. & Christophe, A. (2007). Brain responses in 4-month-old infants are already language specific. *Current Biology*, 17(14), 1208-1211.
- Friederici, A. D., Mueller, J. L. & Oberecker, R. (2011). Precursors to Natural Grammar Learning: Preliminary Evidence from 4-Month-Old Infants. *PLoS ONE*, 6(3): e17920.
doi:10.1371/journal.pone.0017920
- Frota, S. (in press). The intonational phonology of European Portuguese. In Sub-Ah Jun (ed.) *Prosodic Typology II*. Oxford: Oxford University Press, 6-42.

- Frota, S. (2002). Nuclear falls and rises in European Portuguese: A phonological analysis of declarative and question intonation. *Probus*, 14-1 (Special Issue on Intonation in Romance, edited by José- Ignacio Hualde), 113-146.
- Frota, S. (2000). Prosody and focus in European Portuguese. Phonological phrasing and intonation. New York: Garland Publishing.
- Geffen, S. & Mintz, T. (2011). Seven-month-olds discrimination of statements and questions. *Proceedings of the 36th Boston University Conference on Language Development*, edited by Alisa K. Biller, Esther Y. Chung, and Amelia E. Kimball. Somerville, MA: Cascadilla Press.
- Gussenhoven, C. (2007). The phonology of intonation. In Paul de Lacy (ed.), *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, 253-280.
- Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge University Press, Cambridge.
- He, C., Hotson, L. & Trainor, L. J. (2007). Mismatch responses to pitch changes in early infancy. *Journal of Cognitive Neuroscience*, 19(5), 878-92.
- Höhle, B., Bijeljic-Babic, R., Herold, B., Weissenborn, J. & Nazzi, T. (2009). Language specific prosodic preferences during the first half year of life: Evidence from German and French infants. *Infant Behavior and Development*, 32(3), 262-274.
- Hyman, L. M. & Monaka, K. C. (2011). Tonal and Non-Tonal Intonation in Shekgalagari. In S. Frota, G. Elordieta & P. Prieto (Eds.), *Prosodic Categories: Production, Perception and Comprehension*. Studies in Natural Language and Linguistic Theory. Springer, Vol. 82, 267-289.
- Jusczyk, P. W., Cutler, A. & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, 64(3), 675-687.
- Koegel, L. K., Koegal, R. L., Green-Hopkins, I. & Barnes, C. C. (2010). Brief report: Question-asking and collateral language acquisition in children with autism. *Journal of Autism and Developmental Disorders*, 40(4), 509-515.

- Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, 5, 831-843.
- Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S. & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9(2), F13-F21.
- Ladd, D. R. (2008). *Intonational phonology*. Cambridge: Cambridge University Press. 2nd Edition.
- Maiden, M. & Robustelli, C. (2000). *A reference grammar of Modern Italian*. London: Edward Arnold.
- Mateus, M. H. M., Brito, A. M., Duarte, I., Faria, I. H., Frota, S., Matos, G., Oliveira, F., Vigário, M. & Villalva, A. (2003). *Gramática da Língua Portuguesa*. Lisboa: Editorial Caminho.
- Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy*, 10(3), 241-265.
- Mattock, K., Molnar, M., Polka, L., & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition*, 106(3), 1367–1381.
- Mazuka, R., Cao, Y., Dupoux, E. & Christophe, A. (2010). The development of a phonological illusion: a cross-linguistic study with Japanese and French infants. *Developmental Science*, 14(4), 693-699.
- Meints, K. & Woodford, A. (2008). Lincoln Infant Lab Package 1.0: A new programme package for IPL, Preferential Listening, Habituation and Eyetracking [WWW document: Computer software & manual]. URL: <http://www.lincoln.ac.uk/psychology/babylab.htm>
- Narayan, C. R., Werker, J. F. & Beddor, P. S. (2010). The interaction between acoustic salience and language experience in developmental speech perception: evidence from nasal place discrimination. *Developmental Science*, 13(3), 407-420.
- Nazzi, T., Bertoncini, J. & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3), 756-766.
- Nazzi, T., Floccia, C. & Bertoncini, J. (1998). Discrimination of pitch contours by neonates. *Infant*

Behavior & Development, 21(4), 779-784.

Nespor, M. & Vogel, I. (2007). *Prosodic Phonology*. Berlin. Mouton De Gruyter. Originally published in 1986 (Dordrecht. Foris).

Newport, E., Gleitman, H. & Gleitman, L. (1977). Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. In C. Snow & C. Fergusson (eds.), *Talking to Children: Language Input and Acquisition*. New York: Cambridge University Press.

Papoušek, M., Bornstein, M. H., Nuzzo, C., Papoušek, H. & Symmes, D. (1990). Infant responses to prototypical melodic contours in parental speech. *Infant Behavior and Development, 13*, 539-545.

Pierrehumbert, J. & Hirschberg, J. (1990). The meaning of intonation contours in the interpretation of discourse. In P.R. Cohen, J. Morgan & M. Pollack (Eds): *Plans and intentions in communication and discourse*. MIT Press.

Polka, L., Colantonio, C. & Sundara, M. (2001). A cross-language comparison of /d/-/ð/ perception: Evidence for a new developmental pattern. *Journal of the Acoustical Society of America, 109(5)*, 2190-2201.

Prieto, P. & Rigau, G. (2007). The syntax-prosody interface: Catalan interrogative sentences headed by *que*. *Journal of Portuguese Linguistics, 6 (2)*, 29-59.

Rialland, A. (2007). Question prosody – an African perspective. In T. Riad & C. Gussenhoven. *Tones and tunes vol. 1 – Typological studies in word and sentence prosody*. Berlin: Mouton de Gruyter.

Saffran, J. R., Werker, J. F., & Werner, L. A. (2006). The infant's auditory world: Hearing, speech, and the beginnings of language. In D. Kuhn & M. Siegler (Eds.), *The 6th edition of the handbook of child psychology*. New York: Wiley, pp. 58–108

Sansavini, A. (1997). Neonatal perception of the rhythmical structure of speech: The role of stress patterns. *Early Development and Parenting, 6(1)*, 3-13.

Sansavini, A., Bertoncini, J. & Giovanelli, G. (1997). Newborns discriminate the rhythm of

- multisyllabic stressed words. *Developmental Psychology*, 33(1), 3-11.
- Sato, Y., Sogabe, Y. & Mazuka, R. (2009). Development of hemispheric specialization for lexical pitch-accent in Japanese infants. *Journal of Cognitive Neuroscience*, 22(11), 2503-2513.
- Shi, R. (2010). Contextual variability and infants' perception of tonal categories. *Chinese Journal of Phonetics*, 2, 1-9.
- Shukla, M., White, K. & Aslin, R. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proceedings of the National Academy of Sciences*, 108(15), 6038-6043.
- Skoruppa, K., Cristiá, A., Peperkamp, S. & Seidl, A. (2011). English-learning infants' perception of word stress patterns. *Journal of the Acoustical Society of America*, 130, EL50-EL55.
- Skoruppa, K., Pons, F., Bosch, L., Christophe, A., Cabrol, D. & Peperkamp, S. (2013). The development of word stress processing in French and Spanish infants. *Language Learning and Development*, 9(1), 88-104.
- Skoruppa, K., Pons, F., Christophe, A., Bosch, L., Dupoux, E., Sebastián-Gallés, N., Limissuri, R. A. & Peperkamp, S. (2009). Language-specific stress perception by 9-month-old French and Spanish infants. *Developmental Science*, 12 (6), 914-919.
- Snow, C. E. (1972). Mothers' speech to children learning language. *Child Development*, 43, 549-565.
- Soderstrom, M., Ko, E. & Nevzorova, U. (2011). It's a question? Infants attend differently to yes/no questions and declaratives. *Infant Behavior and Development*, 34, 107-110.
- Stager, C. L. & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word learning tasks. *Nature*, 388(6640), 381-382.
- Trehub, S. E., Bull, D. & Thorpe, L. (1984). Infants' perception of melodies: The role of melodic contour. *Child Development*, 55, 821-830.
- Vigário, M. (2003). *The Prosodic Word in European Portuguese*. (Interface Explorations Series, 6). Berlin/ New York: Mouton de Gruyter.
- Vigário, M., Frota, S., Martins, F. & Butler, J. (in progress). Characterizing Child Directed Speech in

European Portuguese: prosodic words and intonational phrases.

http://ww3.fl.ul.pt/laboratoriofonetica/babylab/CDS_EP_Prosoody/CDS_EP_prosodic_words_prosodic_phrases.pdf

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7(1), 49-63.

Wong, W. Y. P., Chan, M. K. M. & Beckman, M. E. (2005). An autosegmental-metrical analysis and prosodic annotation conventions for Cantonese. In S. A. Jun (Ed.), *Prosodic Typology – The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press.

Yeung, H. H., Chen, K. H. & Werker, J. F. (2013). When does native language input affect phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language*, 68(2), 123-139.

Table 1: Acoustic analysis of the stimuli.

	Statements		Questions	t-test
	natural	manipulated		
Pitch fall 1 st syll (Hz)	28	67	66	0.12, p = .9*
F0 range 2 nd syll (Hz)	-25	-25	192	23.46, p<.001
Final F0 (Hz)	163	163	380	23.61, p<.001
Duration (ms)	529	529	765	11.91, p<.001

*the t-test reported is between the manipulated dec and int values, t-test between natural dec and int significantly different, 10.56, p<.001

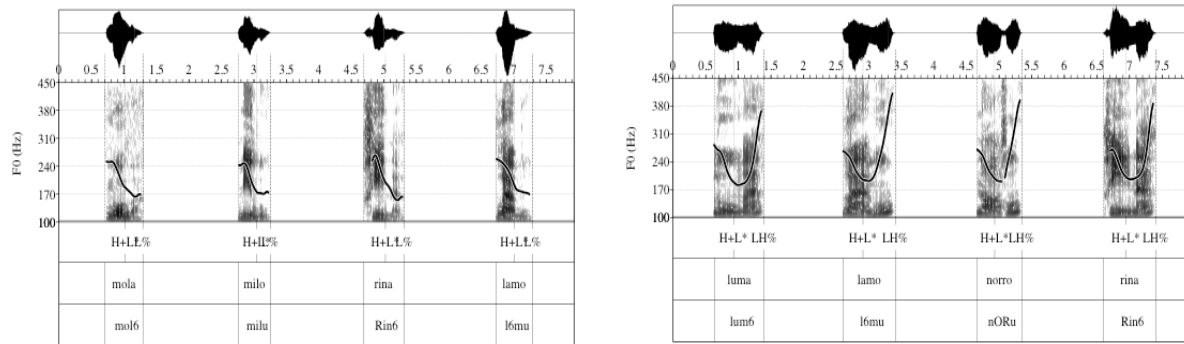


Figure 1: Intonation contours of illustrative statement (left panel) and question (right panel) stimuli, with prosodic labelling and orthographic and phonetic transcription (in SAMPA).

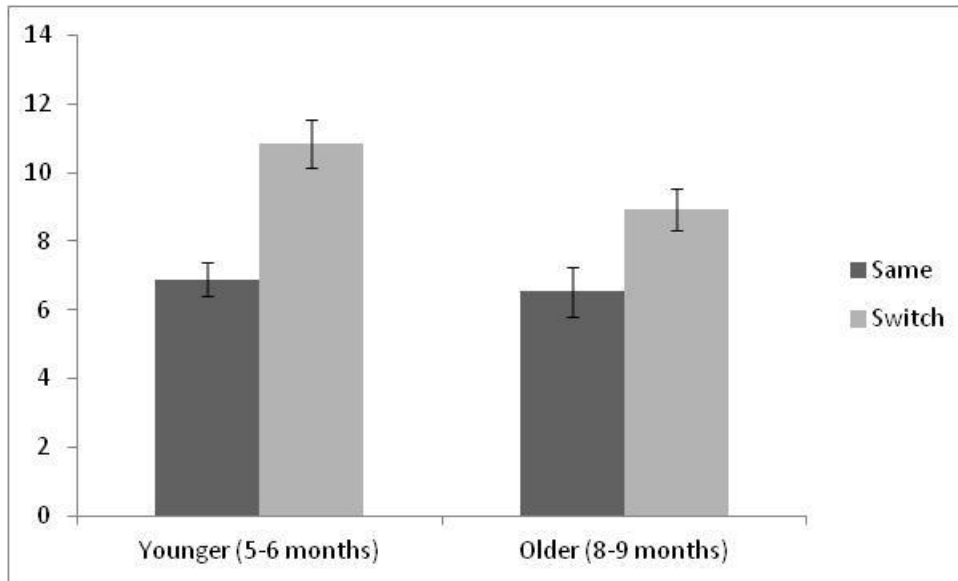


Figure 2: Average looking times (in seconds) to the same/switch test trials, across the two age groups. Error bars indicate the standard error of the mean.