INFANT PERCEPTION OF PROSODIC BOUNDARIES WITHOUT THE PAUSE CUE: AN EYE-TRACKING STUDY

Sónia Frota[1], Joseph Butler[1], Cátia Severino[1], Ertugrul Uysal[2] and Marina Vigário[1]

[1]University of Lisbon, [2]Université de Neuchâtel
sfrota@campus.ul.pt, drjoebutler@gmail.com, catiaseverino@edu.ulisboa.pt, ertugrul.uysal@unine.ch, mvigario@campus.ul.pt

ABSTRACT

Prosodic boundaries play a crucial role in signaling speech chunking, and may thus facilitate language learning. Previous studies have shown that infants are sensitive to prosodic boundaries and use them to segment speech. As prosodic boundary cues vary across languages, infants' sensitivity to prosodic boundaries may also vary. The present study explores the perception of prosodic boundaries without the pause cue in European Portuguese 9-month-old infants. Using a familiarization procedure with visual fixation implemented with eye-tracking, infants were presented with sequences of delexicalized utterances with and without a prosodic boundary while watching a video with a randomly moving pattern. Successful discrimination was found, demonstrating that the pause is not a necessary cue by 9 months in line with the language-specific adult pattern. Potential relations of discrimination abilities with later language outcomes are examined, and implications of our findings for crosslinguistic variation in the development of prosodic boundary perception are discussed.

Keywords: infant perception, intonational phrase boundary, language development, eye-tracking

1. INTRODUCTION

Prosody plays a crucial role in the organization of speech. An utterance consists of prosodic groupings which chunk the speech continuum and are organized into higher level phrases, such as the intonational phrase, and lower level ones, such as the phonological phrase ([7], [20], [26]). Prosodic phrases, in particular the higher level ones, tend to be signalled by acoustic cues like major pitch changes (i.e., pitch lowering or pitch rising), lengthening, or the presence of a pause ([7], [15], [30]).

The prosodic organization of language interfaces with other linguistic domains. For example, the intonational phrase (IP) bears a relation to a clause-like syntactic unit ([10], for a review), and thus sentence or clause boundaries are usually aligned with IP boundaries. Consequently, prosodic boundary cues can be used in language processing to discover aspects of syntactic structure ([5]). Moreover, prosodic boundary cues can be perceived in the absence of identifiable meaning based on words and segmental cues, as in delexicalized sentences or artificial grammar learning ([16], [21]).

Previous studies have shown that infants, during the first year of life, are sensitive to prosodic boundaries. Infants prefer listening to speech with a pause at a natural prosodic boundary (e.g., a clausal boundary), than with a pause that disrupts the prosodic organization of utterances ([12]). Prosodic boundaries have been shown to facilitate infants' initial word segmentation attempts and word learning. Successful segmentation is obtained at utterance edges, which are typically marked by a pause as well as duration and pitch cues, earlier than in utterance-medial positions ([3], [14]). Sound sequences aligned with utterance internal major boundaries are more easily recognized and mapped onto visual referents than when straddling a prosodic boundary ([11], [27]).

Prosodic boundary cues, however, may vary across languages ([7]). For example, in American English a pitch change seems to be a necessary boundary cue, whereas in German both a pitch change and lengthening are necessary, and in Dutch it is the pause cue that is weighed higher ([13], [31]). Infants' sensitivity to prosodic boundary cues seems to attune to the language-particular pattern by 6-8 months ([13], [23], [24], [31]).

The present study investigates the perception of prosodic boundaries in European Portuguese-learning infants. This is the first attempt to examine infants' perception of prosodic boundaries without the pause cue in European Portuguese (EP). The only previous study looking at prosodic boundaries tested word segmentation abilities ([3]), like most studies for other languages ([13], [14], [23], [24]), and focused on utterance boundaries cued by a combination of pitch lowering, final lengthening and pause. Segmentation was found as early as by 4 months at the utterance edge, but was not yet fully developed in utterance-medial position by 10 months. In the current study we tested 9 month-olds' discrimination of utterances with and without an internal IP boundary, which is cued by pitch rising and final lengthening. EP displays an unusual combination of prosodic properties, described as a Romance and Germanic mix, with strong cues to higher level boundaries, namely the utterance and intonational phrase, as well as to word boundaries, but not to lower phrase boundaries ([8], [29]). Adult studies have shown that both pitch change and preboundary lengthening are robust cues to signal higher level boundaries in EP, with the pause not being a necessary cue ([6], [25]). If EP-learning infants are sensitive to higher level boundaries, and attunement to the language-specific pattern of boundary cues is manifested by 6-8 months of age, we expect that 9 month-olds will display discrimination of the contrasting prosodic groupings.

Unlike in previous studies, we did not present infants with speech materials with words or word-like sequences. Delexicalized utterances were used, so that only prosodic information was preserved. Some studies on infant and child processing of IP boundaries using ERPs have claimed that some processing of syntactic information is required together with prosody ([17], [18]). A similar argument could be built around the presence of lexical or word-like information for the processing of IP boundaries. Using delexicalized materials allowed us to discard any possible influence of word combination patterns on our results, or even of the presence of word-like and clause-like structures, along the lines of adult studies ([21]). If evidence for discrimination is found in our study, this can only be due to the processing of prosodic structure.

Another novel feature of the current study is the use of eye-tracking to implement a modified version of the familiarization-preference procedure, which has been successfully applied to the study of language discrimination ([1], [2], [28]). Besides providing looking measures independent of experimenter coding, eye-tracking might be useful to establish more sensitive measures by exploring selective looking to colorful images/patterns while listening to auditory stimuli.

The present study is also the first attempt to explore relations between infants' prosodic boundary discrimination abilities and later language outcomes. Given that higher level prosodic boundaries facilitate word segmentation, and tend to be aligned with major syntactic boundaries, infants may exploit prosodic boundary cues to learn about the lexicon and syntax ([4], [19]). It might thus be expected that infants' discrimination abilities are positively correlated with later language development.

2. METHOD

2.1. Participants

Fifteen typically developing infants from monolingual EP homes participated in this study (7 females, mean age 9 months 10 days, range 8 months 6 days – 10 months 27 days). Five additional infants were rejected (2 for failing to complete the experiment, 1 for living in a bilingual household, 1 for having an older, autistic sibling, and 1 for low looking time, i.e. less than 200 ms, at any of the conditions).

2.2. Stimuli

The stimuli consisted of two pairs of short sentences as in (1), that naturally have two distinct prosodic groupings: the (a) sentences are SVO utterances phrased into one IP; the (b) sentences start with a topic phrase that constitutes an IP on its own, and thus include two IPs.

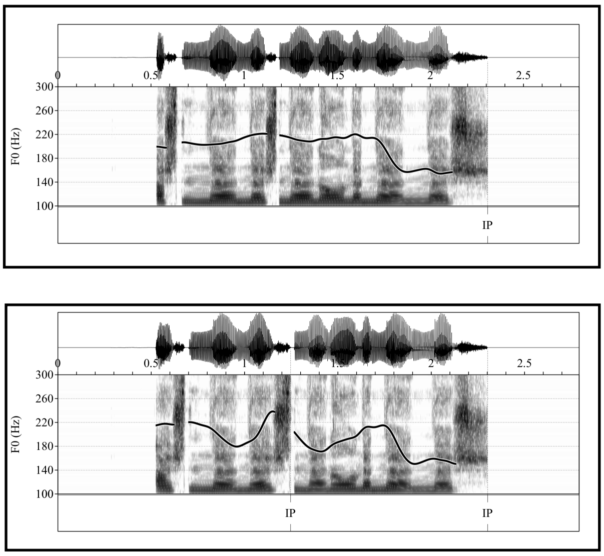
(1) a. (As meninas deram bonecas) $_{\mathbf{IP}}$
'The girls gave dolls'
b. (Às meninas) $_{\mathbf{IP}}$ (deram bonecas) $_{\mathbf{IP}}$
'To the girls, (they) gave dolls'

The sentences were recorded by a female, native EP speaker. Two productions of each sentence were selected. The eight items were then delexicalized using Mbrola, according to [22]. Delexicalization was obtained by converting all vowels into the [ɐ] vowel and all consonants into [n], except for coda consonants that were converted into [ʃ]. All the original prosodic information was preserved. Manipulated versions of the sentences in (1) are illustrated in Fig. 1 (with respective sound files).

The mean length of the sentences with and without the internal IP was 1817ms and 1709ms for sentence 1, and 1570ms and 1531ms for sentence 2. Acoustic analysis of pitch and duration on the syllable before and after the boundary, and on their counterparts in the sentences without boundary, revealed that the two prosodic groupings were indeed distinguished by the presence vs. absence of IP boundary cues. The with-boundary sentences showed preboundary pitch rise (mean 75Hz), preboundary lengthening (mean 283ms), and pitch reset after the boundary (mean 171Hz), unlike the without-boundary ones (mean 14Hz, 190ms, 217Hz, respectively).

The two tokens from each sentence were used to create sound files for presentation as familiarization trials and test trials. Tokens were randomly concatenated with a silent interval of 1500ms between them to produce familiarization files of similar length (~120 s). Test trials consisted of four tokens of the same sentence with a 1500ms silence in between. The length of a test trial was 14000ms.

Figure 1: Examples of delexicalized utterances without (top) and with (bottom) the internal IP.



2.3. Procedure

A modified version of the familiarization-preference procedure was used ([1], [2], [28]), implemented with a SMI RED500 eye-tracker. Infants were seated on a caregiver's lap in front of the eye-tracker's monitor, with speakers hidden behind the monitor. The experiment included two blocks (one for each sentence pair), each consisting of familiarization and test. In the familiarization phase, infants were presented with either the with-IP or without-IP sequences (counterbalanced), while watching a video with a randomly moving pattern with varying shape and size. Familiarization continued until the infant accumulated 60 s of total looking time to the screen or until the end of the familiarization string. The test phase had four trials, two without and two with IP (two familiar, two novel), presented in randomized order with the constraint that the first two test trials were different. Infants listened to the test trials while watching the same video as in the familiarization. At the end of the familiarization phrase and in between test trials, a colorful image was displayed. After a fixation of 400ms to the image, the experiment moves on to the next trial. A video with a congratulatory message ends each of the two blocks.

2.4. Measures of language outcomes

The infants were part of a longitudinal study relating infant perceptual skills to later language outcomes. Infants' caregivers completed the EP version of the CDI short forms ([9]) at 12, 18 and 24 months. The EP-CDI short forms (SF) are a parental checklist measure of the child's vocabulary, and of the ability to combine words generally noted as a significant milestone of syntactic development.

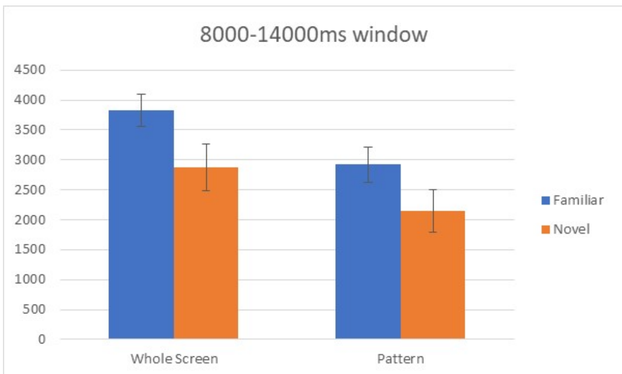
3. RESULTS

3.1. Prosodic boundary discrimination

Two areas of interest were defined – AOI1, the whole screen, and AOI2, a dynamic AOI corresponding to the visual moving pattern – and total looking time to each AOI was extracted along the test trial. An inspection of the proportion of looks to the familiar and novel sequences during the time course of test trials allowed us to find a time window of interest for the effect of familiarity, which corresponded to the second half of the test trial (8000ms-14000ms). Any consistent difference in looking time between familiar and novel is taken as an indication of discrimination abilities (e.g., [2]).

No difference was found in the familiarization looking time between infants familiarized with sequences without-IP and with-IP ($t(13)=.333$, $p=.745$). Looking times to familiar and novel in the time window of interest for the two AOIs are presented in Fig. 2.

Figure 2: Mean looking times (ms) to familiar and novel across the two AOIs. Error bars indicate the standard error of the mean.



A repeated measures ANOVA with the within-subject factor of familiarity (familiar, novel) and the between subject factor of familiarization condition (without-IP, with-IP) was carried out for each of the two AOIs. For AOI1, a main effect of familiarity was found ($F(1,13)=5.536$, $p=.035$, $\eta^2=.299$), but no effect of condition ($F(1,13)=.236$, $p=.635$, $\eta^2=.018$), and no interaction ($F(1,13)=.246$, $p=.628$, $\eta^2=.019$). Similarly, for AOI2 the analysis revealed a significant effect of familiarity ($F(1,13)=5.785$, $p=.032$, $\eta^2=.308$), no effect of condition ($F(1,13)=.024$, $p=.879$, $\eta^2=.002$), and no interaction ($F(1,13)=.010$, $p=.923$, $\eta^2=.001$). However, effects were slightly stronger in AOI2 (with a medium effect size compared to a small effect size in AOI1).

Independent of the familiarization string condition, EP-learning infants listened longer to familiar test trials than to novel ones. This demonstrates evidence for discrimination of the contrasting prosodic groupings.

3.2. Later language outcomes

The potential link between discrimination abilities for prosodic boundaries and later language outcomes was examined by assessing the correlation between looks to familiar minus novel and EP-CDI scores for vocabulary and word combinations. A near-significant correlation was found between discrimination performance at 9 months and the ability to combine words at 24 months ($r=.871$, $p=.055$), suggesting that perception of prosodic boundaries may be related to early development of syntax in production.

We also compared the EP-CDI scores of the infants included in the discrimination experiment with those of the infants excluded. The infants excluded showed lower expressive vocabulary scores at 18 months (mean excluded 3.3, mean included 35.8) and 24 months (mean excluded 36.4, mean included 52.3), suggesting that perception of prosodic boundaries may be related to faster development of the lexicon.

4. DISCUSSION

In this study, we investigated the perception of prosodic boundaries in European Portuguese-learning infants. Using a modified version of the familiarization-preference procedure implemented with eye-tracking, we have shown that 9 month-olds successfully discriminate between utterances with and without an internal IP boundary. Given that no pause cue was present in the sound materials, and the IP boundary was thus only signaled by the most common cues used in EP, namely pitch changes and lengthening, this result is in line with the adult language-specific pattern. It thus provides further support for infants' attunement to the language-particular pattern of boundary cues during the first year of life. EP-learning infants were found to behave similarly to English and German infants ([23], [31]), and differently from Dutch infants, who require the presence of the pause cue ([13]). However, unlike German 8-month-olds that only showed discrimination when familiarized with sequences without an internal IP boundary ([31]), EP infants' discrimination was not affected by the type of prosodic grouping heard during familiarization. Interestingly, the properties of the EP prosodic system are different from those of English, Dutch and German, that are arguably more prosodically similar ([8], [15], [29]). The reasons behind cross-linguistic differences in infants' perception of prosodic boundaries require further research.

Unlike previous studies, this study used delexicalized utterances, thus eliminating any possible effects of the presence of word-like and clause-like structures in the perception of IP boundaries (as suggested by [18] and [19]). Consequently, infants' successful discrimination could only be due to the processing of prosodic structure. This is relevant to prosodic bootstrapping theory, that holds infants can exploit prosodic boundary cues to learn about the lexicon and syntax ([4], [19]).

Having demonstrated infants' ability to perceive prosodic boundaries in the absence of other (nonprosodic) cues to word-like and clause-like structures, the question arises whether this ability relates to later language outcomes. We started to address this question by examining how discrimination performance correlated with CDI scores later in development. The tendency shown in the results suggests that perception of IP boundaries at 9 months may be related to early development syntax and the lexicon. If a similar pattern is found in future studies with larger samples, prosodic boundary perception will offer a potential measure to predict aspects of language development, and to detect language impairment at an early age.

On the methodological side, the use of eye-tracking in the present study was shown to provide not only experimenter independent, but also more accurate (time window) and sensitive (AOIs) measures of discrimination abilities.

5. ACKNOWLEDGEMENTS

This research was supported by the Portuguese Foundation for Science and Technology (Grants EXCL/MHC-LIN/0688/2012 and PTDC/MHCLIN /3901/2014) in conjunction with the European Regional Development Fund from the EU, Portugal 2020 and Lisboa 2020 (Grant PTDC/LLT-LIN/29338/2017).

6. REFERENCES

[1] Bosch, L., Sebastián-Gallés, N. 2001. Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, 2(1), 29-49.

[2] Bosch, L., Sebastián-Gallés, N. 2003. Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech*, 46(2-3), 217-243.

[3] Butler, J., Frota, S. 2018. Emerging word segmentation abilities in European Portuguese-learning infants: New evidence for the rhythmic unit and the edge factor. *Journal of Child Language*, 45, 1294-1308.

[4] Christophe, A., Millotte, S., Bernal, S. Lidz, J. 2008. Bootstrapping lexical and syntactic acquisition. *Language and Speech*, 51(1-2), 61-75.

[5] Cutler, A., Dahan, D., Van Donselaar, W. 1997. Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40(2), 141-201.

[6] Frota, S. 2000. *Prosody and Focus in European Portuguese: Phonological Phrasing and Intonation*. New York: Garland.

[7] Frota, S. 2012. Prosodic structure, constituents and their implementation. In: Cohn, A. C., Fougeron, C., Huffman, M. (eds.), *The Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press, 255–65.

[8] Frota, S. 2014. The intonational phonology of European Portuguese. In: Jun, S.-A. (ed.), *Prosodic typology II: The Phonology of Intonation and Phrasing.* Oxford: Oxford University Press, 6–42.

[9] Frota, S., Butler, J., Correia, S., Severino, C., Vicente, S., Vigário, M. 2016. Infant communicative development assessed with the European Portuguese MacArthur-Bates Communicative Development Inventories Short forms. *First Language*, 36(5), 525-545.

[10] Frota, S., Vigário, M. 2018. Syntax-phonology interface. In: Aronoff, M. (ed.), *Oxford Research Encyclopedia* in Linguistics. Oxford: Oxford University Press, http://linguistics.oxfordre.com/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-111

[11] Gout, A., Christophe, A., Morgan, J.L. 2004. Phonological phrase boundaries constrain lexical access II. Infant data. *Journal of Memory and Language*, 51(4), 548-567.

[12] Hirsh-Pasek, K., Nelson, D.G.K., Jusczyk, P.W., Cassidy, K.W., Druss, B., Kennedy, L. 1987. Clauses are perceptual units for young infants. *Cognition*, 26(3), 269-286.

[13] Johnson, E.K., Seidl, A. 2008. Clause segmentation by 6-month-old infants: A crosslinguistic perspective. *Infancy*, 13(5), 440-455.

[14] Johnson, E.K., Seidl, A., Tyler, M.D. 2014. The edge factor in early word segmentation: Utterance-level prosody enables word form extraction by 6-month-olds. *PloS one*, 9(1), e83546.

[15] Ladd, D.R., 2008. *Intonational Phonology*. Cambridge: Cambridge University Press.

[16] Langus, A., Marchetto, E., Bion, R.A.H., Nespor, M. 2012. Can prosody be used to discover hierarchical structure in continuous speech?. *Journal of Memory and Language*, 66(1), 285-306.

[17] Männel, C., Friederici, A.D. 2011. Intonational phrase structure processing at different stages of syntax acquisition: ERP studies in 2-, 3-, and 6-year-old children. *Developmental Science*, 14(4), 786-798.

[18] Männel, C., Schipke, C.S., Friederici, A.D. 2013. The role of pause as a prosodic boundary marker: Language ERP studies in German 3-and 6-year-olds. *Develop. Cognitive Neuroscience*, 5, 86-94.

[19] Morgan, J. L., Demuth, K. 1996. *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

[20] Nespor, M., Vogel, I. 2007. *Prosodic Phonology: with a new foreword.* Berlin & New York: Walter de Gruyter.

[21] Pannekamp, A., Toepel, U., Alter, K., Hahne, A., Friederici, A.D. 2005. Prosody-driven sentence processing: An event-related brain potential study. *Journal of Cognitive Neuroscience*, 17(3), 407-421.

[22] Ramus, F., Mehler, J. 1999. Language identification with suprasegmental cues: A study based on speech resynthesis. *J. Acoust. Soc. Am.* 105(1), 512-521.

[23] Seidl, A. 2007. Infants' use and weighting of prosodic cues in clause segmentation. *Journal of Memory and Language*, 57(1), 24-48.

[24] Seidl, A., Cristià, A. 2008. Developmental changes in the weighting of prosodic cues. *Developmental Science*, 11(4), 596-606.

[25] Severino, C.S., 2016. Perception of phrasal prosody in the acquisition of European Portuguese. (Doctoral dissertation). Lisboa: Universidade de Lisboa.

[26] Shattuck-Hufnagel, S., Turk, A.E. 1996. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, 25(2), 193-247.

[27] Shukla, M., White, K.S., Aslin, R.N. 2011. Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *PNAS*, 108(15), 6038-6043.

[28] 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.

[29] Vigário, M. 2003. *The Prosodic Word In European Portuguese*. Berlin & New York: Mouton de Gruyter.

[30] Wagner, M., Watson, D.G. 2010. Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes*, 25(7-9), 905-945.

[31] Wellmann, C., Holzgrefe, J., Truckenbrodt, H., Wartenburger, I., Höhle, B. 2012. How each prosodic boundary cue matters: Evidence from German infants. *Frontiers in Psychology*, 3, 580.