10th International Conference on Speech Prosody 2020
25-28 May 2020, Tokyo, Japan

Looking for the edge: Emerging segmentation abilities in atypical development

Sónia Frota, Jovana Pejović, Cátia Severino, Marina Vigário

Center of Linguistics, University of Lisbon, Portugal
sfrota@campus.ul.pt, jpejovic@edu.ulisboa.pt,
catiaseverino@campus.ul.pt, mvigario@campus.ul.pt

Abstract

The ability to extract word-forms from continuous speech plays a crucial role in language acquisition, particularly for vocabulary development. This ability develops differently across languages, and was recently shown to be modulated by prosody: words at prosodic edges are segmented earlier than in utterance-medial position. Studies on word segmentation in atypical development are scarce, and the role of prosody has not been explored. The current study investigated emerging segmentation abilities in European Portuguese (EP)-learning infants with Down Syndrome (DS), and in a mixed group of infants at-risk for language impairment (AR). We examined whether prosody facilitated word segmentation, as previously shown for EP typically developing infants (TD), and whether segmentation abilities correlated with concurrent language skills. DS, unlike AR infants, did not segment either at prosodic edge or medial position. AR, like TD, successfully segmented at the edge only, but unlike TD did not show the emergence of segmentation in medial position. Both DS and AR infants' segmentation abilities at the prosodic edge were correlated with concurrent vocabulary scores measured with the CDI. Our findings demonstrate that prosody drives early segmentation abilities, and that segmentation abilities and receptive and expressive language skills are closely linked in atypical development.

Index Terms: early word segmentation, prosodic edge, atypical development, Down Syndrome, infants at-risk for language impairment, language acquisition

1. Introduction

Word segmentation is the ability to extract word-forms from continuous speech. This ability plays a central role in language acquisition, particularly for word learning, as word segmentation in the first year of life has been associated with later lexical skills ([1]-[5]). Although current knowledge suggests that segmentation is a language general ability, segmentation skills are known to develop differently across languages (e.g., [6]; [7], for a review). Furthermore, recent work has shown that segmentation abilities are modulated by prosodic structure: namely, words at prosodic edges are segmented earlier than in utterance-medial position ([7], [8]).

Studies on word segmentation in atypical development are scarce. Three studies have addressed the development of this ability in English-learning toddlers with Williams Syndrome ([9] on 15 to 48 month-olds) and Down Syndrome ([10], [11] on 18-20 month-olds). In the three studies toddlers were familiarized with isolated words and tested with passages. These studies concluded that the development of segmentation was seriously delayed, but the learning path was similar to that found for typical development babies. Moreover, neither age nor concurrent lexical development were found to be related with segmentation skills, and these were not predictive of later language outcomes.

Two other studies reported mixed findings regarding segmentation skills in a group of infants at-risk for language impairments, namely pre-term infants. In a familiarization paradigm with the passages-first order, preterm 8-month-old (maturational age) Spanish or Catalan monolinguals failed to segment CVC or CCVC target words, unlike full-term 8-month-olds ([12]). However, also using a passages-first familiarization experiment, in [13] preterm 6-month-old (post-natal age) French-learning infants were shown to successfully segment CV words, like full-term French infants of the same (post-natal and maturational) age.

While the few previous studies started to examine segmentation skills in atypical development, in none of them the role of prosody has been explored. The current study investigated early segmentation abilities in European Portuguese (EP)-learning infants with Down Syndrome (DS), and in a mixed group of infants at-risk for language impairment (AR), including premature birth and familial risk for autism or language disorder. Our main goal was to examine whether prosody facilitated word segmentation, as previously shown for typically developing infants (TD) in the only study that addressed emerging segmentation abilities in EP-learning infants ([7]). Using a variation of the visual familiarization paradigm, and the passages-first order, TD infants aged 4-10 months old were tested with target words located at utterance-edges and utterance-medial positions. Evidence for segmentation was found as early as 4 months, but only for utterance-edge. However, by 10 months, evidence of the emergence of segmentation in utterance-medial position was found, showing that segmentation abilities are crucially modulated by utterance-level prosody in EP. If similar mechanisms and/or trajectories guide word segmentation in atypical development, a similar pattern of results is expected, albeit possibly delayed.

Furthermore, lexical acquisition has been shown to be delayed in these populations ([14], [15]). Given the link between segmentation abilities and lexical skills in typical development, another goal of the present study was to examine whether segmentation abilities in atypical development correlate with concurrent language skills. Specifically, we asked whether prosody modulated the relation between segmentation abilities and lexical knowledge.

2. Method

2.1. Participants

Two groups of babies participated in this study: 25 infants and toddlers with DS (12 girls, mean age 16 months and 12 days,

10.21437/SpeechProsody.2020-166

range 7 to 23 months), and 21 AR infants and toddlers (10 girls, mean age 15 months and 17 days, range 6 to 26 months). The AR group included a mix of risk factors: preterm birth (10), familial risk for autism or language disorder (9), and other factors like low Apgar score and reanimation at birth (2). Six additional infants did not complete the experiment (4 DS, 2 AR). All participants were raised in monolingual EP homes, had normal hearing to mild hearing loss and normal or corrected-to-normal vision (according to clinical screening). They had no reported history of seizures, neurological sequelae or other serious medical or neurological conditions. Preterm babies had a gestational age below 37 weeks. Babies with familial risk for neurodevelopmental disorders had an older sibling or first degree relative diagnosed with such a disorder. Parents gave written informed consent before participation in the study. The study was approved by the Ethical Committee for Research of the School of Arts and Humanities of the University of Lisbon.

2.2. Stimuli

The same materials as in [7] were used. Target word-forms were four monosyllabic CVC or CVG pseudo-words with a frequent sound pattern in the language. Pseudo-words were embedded in sentences in two prosodic conditions: utterance-edge-final position and utterance medial position (aligned with a lower phrase boundary in half of the cases and to a phrase-internal word in the other half), as illustrated in (1). For each pseudo-word, two six-sentence passages were created, one for each prosodic condition. Mean sentence length was 10 syllables (range 9-11 syllables). Within sentences there were no internal intonational phrase boundaries. Thus, the only major prosodic boundary was that found between sentences. Word lists for each pseudo-word were created from different spoken exemplars, all with different intonation. Word lists included 15 exemplars of each pseudo-word, with a 500 ms pause between each exemplar. Both the passages and word lists were recorded by a female, native EP-speaker, in child-directed speech. The sound stimuli are available at "http://labfon.letras.ulisboa.pt/babylab/infant_word_segmentation/word_segmentation_supporting_materials.htm".

(1) Gostei daquela imagem do sau (major prosodic edge)
'I liked that image of sau'
A Maria tomou sau com limão (lower phrase)
'(The) Mary took sau with lemon'
O Marco deu um sau grande ao Zé (phrase-internal)
'(The) Mark gave a sau big to Joe'

As described in [7], the main acoustic cues distinguishing the two contrasting prosodic conditions are pitch range and duration, with a pitch fall (due to the nuclear contour H+L* L%) and pre-boundary lengthening characterizing the prosodic edge condition (pitch range, mean 60Hz; target word duration, mean 495ms). Utterance medial position (whether lower phrase or phrase internal) shows no large pitch movement (mean 25Hz) or lengthening (mean 309ms).

2.3. Procedure

The procedure was a modified version of the visual familiarization paradigm, identical to that used in [7]. Participants were seated on a caregiver's lap in front of a computer monitor, with speakers hidden behind the monitor. An attractive attention-getting image initiated the experiment. After 2 consecutive seconds of fixation of that image, a trial began. Each trial consisted of a red and black checkerboard paired with a sound file, that continued playing until the child looked away for more than 2 seconds, or the sound file ended. The attractive image was then presented again. All trials presented the same pattern.

Each participant was familiarized with two passages, and tested with two familiar and two unfamiliar word-forms. In the familiarization phase, the two passages (one with the target pseudo-word in utterance-edge-final position, and the other with a different pseudo-word in utterance-medial position) were presented alternatively until 25 seconds of looking time were accumulated to each passage. In the test phase, each of the four word-form lists was randomly presented three times, split into three blocks of four test trials. When all 12 test trials were presented, the experiment ended. The four pseudo-words were counterbalanced so that half of the participants heard two of such pseudo-words as targets (familiar word-forms) and the other two as unfamiliar items. Moreover, within these groups prosodic condition of a given pseudo-word was also counterbalanced. The LOOK software ([16]) was used to control stimuli presentation and to record participants' looking time to the screen. Any consistent difference in looking times to familiar and unfamiliar word-forms is taken as an indication of segmentation abilities.

2.4. Measures of concurrent language skills

Concurrent language skills were assessed with the EP version of the MacArthur Bates Communicative Development Inventory (CDI) short forms ([17]). The EP-CDI short forms are a parental checklist that provides measures of receptive and expressive vocabulary (short form 1, administered up to 18 months), and of expressive vocabulary and the ability to combine words (short form 2, administered between 18 and 30 months). Participants' caregivers completed the EP-CDI short forms at the moment of testing or within two-weeks before or after the visit to the lab.

3. Results

Mean looking times for the familiar word-forms in each prosodic condition (edge and medial), as well as the unfamiliar word-forms, were calculated for the two groups of participants (DS and AR), and compared with the data for TD infants from [7] (Figure 1). First, we inspected word segmentation abilities in each of the two atypical developing groups. A cross-group analysis, also including TD infants, concludes this section.

3.1. Infants and toddlers with Down Syndrome (DS)

A repeated-measures ANOVA with mean looking time as the dependent measure and the within-participant factor of condition (edge, medial, unfamiliar) revealed no effect of condition ($F(2,48)=0.8$, $p=.45$, $\eta^2=.03$). The results thus indicate that overall the DS group is not segmenting the word-forms in none of the prosodic conditions (Figure 1, left).

Given the wide age range of the DS group, we explored whether age influenced performance in the segmentation task. Two age groups were considered, following the CDI age ranges: younger (below 18 months: N=12, mean age 12 months) and older (above 18 months: N=13, mean age 20 months; Figure 2). A repeated-measures ANOVA with the main factors of condition and age group found no effect of

condition ($F(2,46)=0.8$, $p=.41$, $\eta^2=.02$), or age group ($F(1,23)=0.7$, $p=.39$, $\eta^2=.02$), and no interaction ($F(2,46)=0.7$, $p=.46$, $\eta^2=.01$).

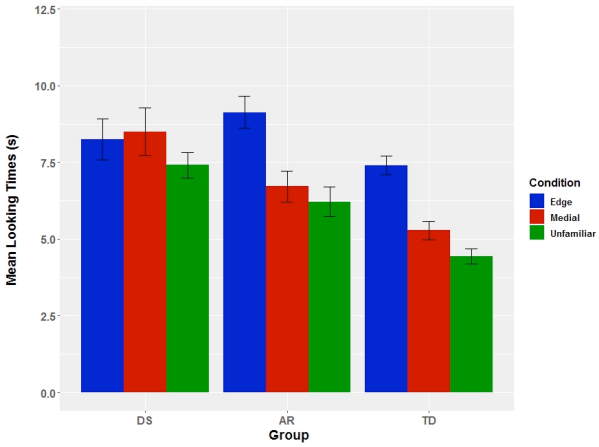

Figure 1: *Mean looking times (s) for the three experimental conditions by group: Down Syndrome (DS), at-risk (AR), and typical developing (TD). Error bars represent standard error of the mean (+/-1).*

Interestingly, in the younger group the difference in looking times between medial and unfamiliar was larger than between edge and unfamiliar, whereas in the older group the reverse pattern obtains (Figure 2). Paired t-tests comparing the three conditions to each other within age groups (corrected p value of .01) showed that only edge and unfamiliar in the older age group approached significance ($p=.08$).

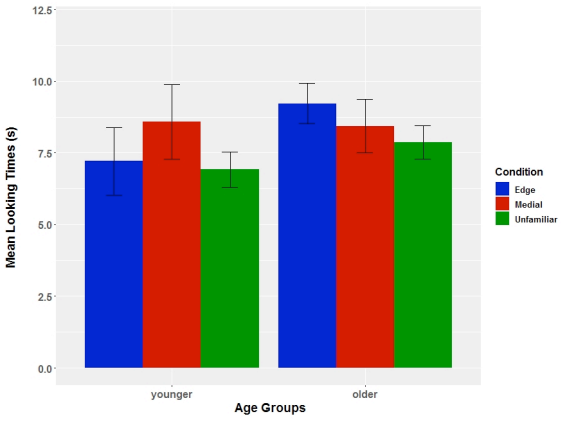

Figure 2: *Mean looking times (s) for the three experimental conditions in the DS group by age (younger, older). Error bars represent standard error of the mean (+/-1).*

Finally, we examined whether the results in the segmentation task correlated with concurrent language skills. Correlation analyses were performed between edge, medial, edge over unfamiliar, medial over unfamiliar (obtained by subtracting unfamiliar from edge or medial), and CDI scores for receptive and expressive vocabulary. A significant positive correlation was found between edge over unfamiliar and the expressive vocabulary score ($r=.34$, $p=.05$). No other correlations were significant ($p>.1$).

3.2. At-risk infants and toddlers (AR)

Unlike the DS group, a repeated-measures ANOVA in the AR group with mean looking time as the dependent measure and the within-participant factor of condition (edge, medial, unfamiliar) revealed a significant effect condition ($F(2,40)=9.64$, $p<.001$, $\eta^2=.32$). Post hoc tests (corrected p value of .01) showed that the effect was driven by the prosodic edge, as there were significant differences between edge and unfamiliar ($p=.008$), but not medial and unfamiliar ($p=.99$), and edge and medial approached significance ($p=.03$). This demonstrates evidence for segmentation in utterance-edge-final position only (cf. Figure 1, mid).

To examine whether age influenced performance in the AR risk group, the group was split into younger (N=13, mean age 11 months) and older (N=8, mean age 21 months), as shown in Figure 3. A repeated-measures ANOVA with the main factors of condition and age group found a main effect of condition ($F(2,38)=11.26$, $p<.001$, $\eta^2=.37$), no effect of age group ($F(1,19)=1.69$, $p=.2$, $\eta^2=.08$), and no interaction ($F(2,38)=1.74$, $p=18$, $\eta^2=.08$). Paired t-tests comparing the three conditions to each other within age groups (corrected p value of .01) approached significance for edge and unfamiliar ($p=.03$) and edge and medial ($p=.08$) in the younger group, and were significant in the older group (respectively, $p=.007$ and $p=.01$). Thus, the same pattern of results was found across age groups.

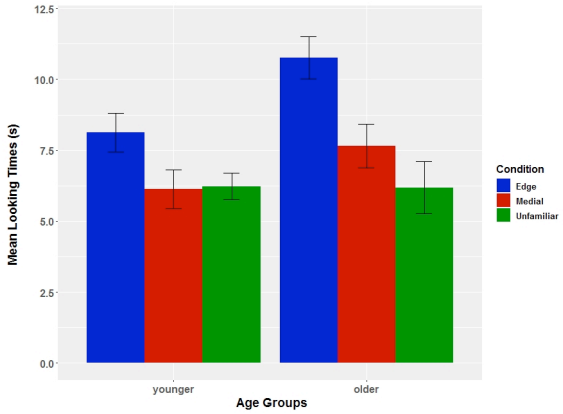

Figure 3: *Mean looking times (s) for the three experimental conditions in the AR group by age (younger, older). Error bars represent standard error of the mean (+/-1).*

Like for DS, correlation analyses were performed between the word segmentation measures and CDI measures. A significant positive correlation was found between edge and receptive vocabulary ($r=.56$, $p=.02$). No other correlations were significant.

3.3. Cross-group analysis: DS, AR and TD

Although the separate analyses for DS and AR suggested that the development of segmentation abilities differs in the two groups, it is unknown whether the two groups are indeed significantly different, and whether they deviate from the pattern that characterizes emerging segmentation abilities in TD infants (cf. Introduction, and Figure 1, right). To address these questions, a cross-group analysis was performed on the

DS and AR data from the current study and the TD data (N=40) from [7]. Given the unbalanced sample sizes across groups, a mixed-model analysis in the *lmerTest* package in R ([18], [19]) was computed with looking time as the dependent measure. Fixed factors were condition and group (the model allowed the interaction) and random effect was *by-subject intercept*.

Type III ANOVA with Satterthwhite's method revealed a main effect of condition ($F(2,172)=18.5$, $p<.001$), group ($F(2,86)=8.1$, $p<.001$), and a significant interaction condition *group ($F(4,172)=2.9$, $p=.02$). To ascertain what was driving the interaction, a posthoc pairwise analysis using *emmeans* package in R ([20]) was performed (Bonferroni-corrected). It was observed that the difference between edge and unfamiliar was larger for the TD over DS group (estimate=2.74, SE=.86, $t=2.44$, $p=.01$; Intercept estimate=7.0396, SE=.26, $t=26.255$, $p<.001$), but smaller for the DS over AR group (estimate=-2.06, SE=1.01, $t=-2.06$, $p=.04$), while TD and AR did not differ (estimate=.04, $t=.054$, $p=.95$). Furthermore, the difference between edge and medial was larger for the TD over DS group (estimate=2.37, SE=.86, $t=2.74$, $p=.006$), but smaller for the DS over AR group (estimate=-2.66, SE=1.01, $t=-2.65$, $p= .009$), while TD and AR did not differ (estimate=-.29, t=-.32, p=.75). Finally, the difference between medial and unfamiliar did not differ across the groups (all ts <1, all ps >.5). These findings are depicted in Figure 4.

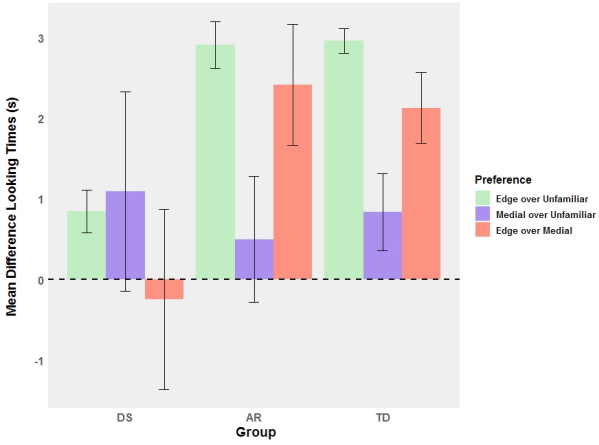

Figure 4: *Mean difference looking times (s) for edge over unfamiliar, medial over unfamiliar and edge over medial, across groups. Error bars represent standard error of the mean (+/-1).*

These results show that the DS group differs from both the AR and TD groups due to shorter looking times to edge relative to unfamiliar, and longer looking times to medial relative to edge. In other words, the DS group is not displaying segmentation at the prosodic edge, while showing a tendency to look longer to medial than the other groups. By contrast, in both TD and AR evidence for segmentation at the prosodic edge was clearly found. Although TD, as a group, show longer looking times to medial relative to unfamiliar than the AR group, reflecting the emergence of segmentation in utterance-medial position for the older infants (described in [7]), this difference did not reach significance. However, considering that the older TD infants were 11 months old and the mean age for the AR group is 15 months, this suggests that utterance-medial segmentation is delayed in the AR group.

4. Discussion

In this study, we have begun investigating early segmentation abilities in European Portuguese-learning infants and toddlers with Down Syndrome (DS), and in a mixed group of infants at-risk for language impairment (AR), including premature birth and familial risk for autism or language disorder. Our main goal was to examine whether prosody facilitated word segmentation, as previously demonstrated for typically developing EP infants (TD), who successfully segmented monosyllabic (CVC/CVG) word-forms located at utterance-edge-final position from 4 months, with utterance-medial segmentation emerging by 10 months ([7]).

DS babies, unlike AR, did not segment either at prosodic edge or medial position. Surprisingly, the younger group of DS babies had a larger difference in looking times between medial and unfamiliar than between edge and unfamiliar. Although not yet segmenting, in the older group of DS babies the difference between edge and unfamiliar approached significance, showing that after 18 months DS babies start looking for the prosodic edge. These findings indicate that, contrary to previous suggestions for English-learning babies with DS ([10], [11]), word segmentation abilities are not only seriously delayed in this group, but they follow a different developmental path. Unlike DS babies, AR babies successfully segmented at the prosodic edge only, following the TD pattern. This study thus provided the first evidence of CVC and CVG segmentation for AR babies (in [12], Catalan and Spanish infants failed). However, unlike TD, AR babies did not show the emergence of segmentation in medial position. This suggests that, although following the TD pattern, segmentation abilities are delayed in this group.

Importantly, segmentation abilities at the prosodic edge were found to be correlated with concurrent vocabulary scores measured with the CDI, both for DS and AR. While a link between segmentation skills and language acquisition has been reported in typical development ([1]-[5]), a similar relation had not yet been found for atypical development ([9], [11], [21]). The finding of such a link has implications for a better understanding of language acquisition trajectories in atypical development, as well as for remediation and intervention strategies.

The few previous studies on word segmentation in atypical development had not considered the role of prosody. Our findings demonstrated that early segmentation abilities are crucially modulated by utterance-level prosody, also in atypical development, and that the ability to segment at the prosodic edge is linked to receptive and expressive language skills in infants and toddlers with Down Syndrome or at-risk for language impairment. Future studies will have to address possible differences due to degree of cognitive impairment in DS, as well as subgroup of at-risk babies.

5. Acknowledgements

This research was supported by the Portuguese Foundation for Science and Technology (Grant PTDC/MHCLIN/3901/2014) in conjunction with the European Regional Development Fund from the EU, Portugal 2020 and Lisboa 2020 (Grant PTDC/LLTLIN/29338/2017). We gratefully acknowledge the collaboration of Diferenças and CEBC-Hospital D. Estefânia in recruitment of participants.

6. References

[1] R. Newman, N. B. Ratner, A. M. Jusczyk, P. W. Jusczyk, and K. A. Dow, "Infants' early ability to segment the conversational speech signal predicts later language development: A retrospective analysis," *Dev. Psychol.*, vol. 42, no. 4, pp. 643–655, 2006.

[2] L. Singh, J. Steven Reznick, and L. Xuehua, "Infant word segmentation and childhood vocabulary development: A longitudinal analysis," *Dev. Sci.*, vol. 15, no. 4, pp. 482–495, 2012.

[3] E. Kidd, C. Junge, T. Spokes, L. Morrison, and A. Cutler, "Individual Differences in Infant Speech Segmentation: Achieving the Lexical Shift," *Infancy*, vol. 23, no. 6, pp. 770–794, 2018.

[4] K. Von Holzen, L.-L. Nishibayashi, and T. Nazzi, "Consonant and Vowel Processing in Word Form Segmentation: An Infant ERP Study," *Brain Sci.*, vol. 8, no. 2, p. 24, 2018.

[5] M. Hoareau, H. H. Yeung, and T. Nazzi, "Infants' statistical word segmentation in an artificial language is linked to both parental speech input and reported production abilities," *Dev. Sci.*, vol. 22, no. 4, p. e12803, 2019.

[6] K. Mersad, L. Goyet, and T. Nazzi, "Cross-linguistic differences in early word form segmentation: a rhythmic-based account," *J. Port. Linguist.*, vol. 10, no. 1, pp. 37-65, 2010.

[7] J. Butler and S. Frota, "Emerging word segmentation abilities in European Portuguese-learning infants: New evidence for the rhythmic unit and the edge factor," *J. Child Lang.*, vol. 45, no. 6, pp. 1294–1308, 2018.

[8] E. K. Johnson, A. Seidl, and M. D. Tyler, "The edge factor in early word segmentation: Utterance-level prosody enables word form extraction by 6-month-olds," *PLoS One*, vol. 9, no. 1, p. e83546, Jan. 2014.

[9] T. Nazzi, S. Paterson, and A. Karmiloff-Smith, "Early word segmentation by infants and toddlers with Williams syndrome," *Infancy*, 2003.

[10] E. Mason-Apps, V. Stojanovik, and C. Houston-Price, "Early word segmentation in typically developing infants and infants with Down Syndrome: A preliminary study," *Proc. 17th Int. Congr. Phonetic Sci. (ICPhS 2011)*, no. August, pp. 1334–1337, 2011.

[11] E. Mason-Apps, V. Stojanovik, C. Houston-Price, and S. Buckley, "Longitudinal predictors of early language in infants with Down syndrome: A preliminary study," *Res. Dev. Disabil.*, vol. 81, pp. 37–51, 2018.

[12] L. Bosch, "Precursors to language in preterm infants. Speech perception abilities in the first year of life.," in O. Braddick, J. Atkinson, and G. M. Innocenti (Eds.), *Progress in Brain Research*, (pp. 239–257). Burlington: Academic Press, 2011.

[13] E. Berdasco-Muñoz, L.-L. Nishibayashi, O. Baud, V. Biran, and T. Nazzi, "Early Segmentation Abilities in Preterm Infants," *Infancy*, vol. 23, no. 2, pp. 268–287, 2018.

[14] A. Sansavini, A. Guarini, and M. C. Caselli, "Preterm birth: Neuropsychological profiles and atypical developmental pathways" *Developmental Disabilities Research Reviews*, vol. 17, pp. 102–113, 2011.

[15] A. McDuffie and L. Abbeduto, "Language disorders in children with mental retardation of genetic origin: Down syndrome, Fragile X syndrome, and Williams syndrome," in R. G. Schwartz (Ed.) *Handbook of Child Language Disorders* (pp. 44–66). Hove: Taylor & Francis, 2008.

[16] K. Meints and A. Woodford, "Lincoln infant lab package 1.0: a new programme package for IPL preferential listening, habituation and eyetracking." 2005. Retrieved from <https://www.lincoln.ac.uk/home/psychology/research/lincolnbabylab/>

[17] S. Frota, J. Butler, S. Correia, C. Severino, S. Vicente, and M. Vigário, "Infant communicative development assessed with the European Portuguese MacArthur–Bates Communicative Development Inventories short forms," *First Lang.*, vol. 36, no. 5, pp. 525–545, 2016.

[18] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest Package: Tests in Linear Mixed Effects Models," *J. Stat. Softw.*, vol. 82, pp. 1–26, 2017.

[19] R Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/, 2019.

[20] R. Lenth, "emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.4.3.01." 2019.

[21] V. Stojanovik, "Prosodic deficits in children with Down syndrome," *J. Neurolinguistics*, vol. 24, no. 2, pp. 145–155, 2011.