A LONGITUDINAL STUDY OF PAUSES, INTERPAUSAL UNITS AND CLAUSES IN INFANT-DIRECTED SPEECH

Anna Kohári and Katalin Mády

Hungarian Research Centre
kohari.anna@nytud.hu, mady.katalin@nytud.hu

ABSTRACT

Infants are able to perceive prosodic features, including pauses, and to detect phrase and clause boundaries from fluent speech. Yet, studies investigating the pauses in infant-directed speech (IDS) with respect to syntactic or prosodic structure are relatively sparse and mostly limited to a given age of the child. Here we analyze 14 mothers' semi-spontaneous narratives in Hungarian IDS compared to adult-directed speech (ADS). The longitudinal analysis showed that the adults tended to talk slower to their 6 and 18 month-old children than to other adults, and that pauses were more frequent in IDS as compared to ADS. Furthermore, the ratio between the frequency of pauses at clause boundaries and within clauses was found to be higher in IDS than in ADS. We also found that in speech directed to 18-month-old infants the interpausal units and clauses tended to consist of less syllables than in ADS.

Keywords: infant-directed speech, pausing, longitudinal study

1. INTRODUCTION

Infant-directed speech (IDS) is characterized by lexical and syntactic simplification, repetition, higher fundamental frequency, and slower speech rate [1] compared to adult-directed speech (ADS). Besides these widely known characteristics, there has been little discussion on pauses in IDS and their relation to syntactic or prosodic structure, although perception tests revealed that infants preferred listening to speech where pauses occur between clauses instead of within clauses [2, 3]. Since the occurrence of pauses could provide crucial cues for infants to perceive clause and phrase boundaries, facilitating the acquisition of linguistic structures [3], in the present work we investigate the complex interplay between the acoustic features (articulation rate, pauses, and length of interpausal units (IPUs)) related to pausing and the clause boundaries.

Although numerous studies have demonstrated that adults talk slower to infants [1, 4, 5, 6, 7], others did not find speech or articulation rate differences between the two registers [8, 9]. A plausible explanation behind these seemingly contradictory results may be that the children's age influences the tempo of IDS. This assumption is supported by the findings of Narayan and McDermott [6], who observed that caregivers gradually increased their speech rate as their infant addressees developed, and by the second year of the children's age their IDS rate approximated that of their ADS.

Several studies have established that pauses tend to last longer in spontaneous IDS than in ADS [1, 5, 10, 11]. However, in an experiment where the participants were asked to describe given objects in a semi-spontaneous manner to adults and to infants, the pause durations in the two registers did not differ significantly [12]. Pause durations in the context of syntactic or prosodic structure received less attention, yet, it has been demonstrated for both registers that the pauses at utterance, clause, or phrase boundaries are typically longer than those within speech units [13, 14, 15]. The two registers were reported to differ in terms of pause durations between independent clauses, which were found to last longer in IDS than in ADS [14]. Furthermore, the aforementioned difference between pause durations at phrase boundaries and at non-boundaries was more pronounced in IDS than in ADS [15].

The available data on the age-related differences of pause durations is also quite sparse. Stern [10] concluded that pauses tended to last longer in American English speech directed towards newborns than to 4-, 12-, and 18-month-old infants, however, no such trend was found for later ages of the children. Kondaurova [14] also found that pauses between clauses tend to be longer when addressing 6-month-olds than 12-month-olds.

Not only the duration, but also the frequency of pauses may be different between the two registers. Several studies concluded that pauses occurred more often in IDS than in ADS [7, 12]. It has also been reported that above 90% of the pauses within IDS were detected between utterances, and only a small fraction in them. This ratio is in huge contrast with the result from ADS, where only 50–68% of the pauses occurred in inter-utterance positions [1, 4, 16].

In our longitudinal study, we address the pausing- and timing-related peculiarities of IDS at different ages of the infants and analyze the extent to which the occurrence of pauses coincides with clause boundaries.

2. MATERIALS AND METHODS

2.1. Participants

Recordings from 14 primiparous mothers (i.e., mothers giving birth for the first time), native speakers of Hungarian were used for our analysis. The ages of the participants ranged from 24 to 41 years (M = 31.6, SD = 5.7), and all of them earned a high school diploma or college degree. The properties of IDS change with the child's age [6], therefore besides investigating the mothers' speech to their newborns, we also repeated the experiments at 6 months of the babies' age when they reached the preverbal stage, and later when the children were 18 month old and could communicate actively.

2.2. Materials

The speakers were asked to tell a fairytale about pixies based on a story book with a sequence of pictures, using their own words, first to an (adult) experimenter (i.e., producing ADS) and then to their own child (IDS). During the ADS recording sessions another experimenter was playing with the children in the same room. Besides the pictures, the pages also contained certain scripted sentences, which the mothers had to incorporate to their otherwise spontaneous storytelling to ensure that the content of the speech was approximately identical in both registers. For the analyses of the present study only the spontaneous speech parts and their pauses were investigated. The first recording session with each mother was conducted in a silent separated room at the Birth Centre of the Military Hospital in Budapest a few days after the childbirth. The experiments were later repeated in the baby lab of the Research Centre for Natural Sciences.

2.2. Analysis

Recordings were annotated [17], and the boundaries of interpausal units and clauses were labelled manually using Praat 6.1.08 [18]. For our analysis a clause was defined as a unit containing a subject and a related predicate with their modifiers, while an incomplete clause lacking a predicate was merged with the nearest and semantically connected complete clause, and the two were then considered as one unit. Articulation rates were acquired by measuring the duration of IPUs and clauses (in seconds) and dividing their syllable counts by the obtained durations. The minimum duration above which a pause was considered for the analysis was set to 30 ms, the typical duration of certain speech sounds in spontaneous speech [19]. The duration of the closure phases of voiceless plosives and affricates occurring before pauses was taken to be 50 ms [19]. Since pauses shorter than 300 ms are traditionally ignored in the IDS-related literature [1, 5, 10], we also repeated the analysis of the investigated parameters, setting the lower limit for the duration of the considered pauses to 300 ms. However, this modification did not affect our findings, except for a single aspect: when only the pauses longer than 300 ms were considered, pause durations within the clauses did not exhibit detectable dependence on the infant's age (whereas applying the 30 ms pause limit they did, see Results). Beside the durations of the pauses their frequencies were also calculated. On the one hand, we determined the number of pauses (normalized to 100 syllables) in the speech of each mother in both registers as a function of the infant's age. On the other hand, the frequencies of pauses at clause boundaries and within clauses were evaluated separately (for each mother, infant age, and register). Since disfluencies may also affect pause durations and articulation rates, and as disfluencies are more common in ADS than in IDS [20, 21, 22], our analysis was restricted to clauses – and the pauses therein – lacking disfluencies (false starts, hesitations). The ratio of pauses between clauses was found to be unaffected by whether we included the disfluencies in the analysis, or omitted them. 4227 IPUs and 3301 clauses were investigated.

Linear mixed effect models were fitted for the dependent variables (i.e., articulation rate; the logarithm of the pause durations) using the R 3.6.2 software [23] with the *lmerTest* package [24]. For the syllable counts, a negative binomial regression mixed model was applied. Register, age, and their interactions were modeled as fixed factors, and subjects as random factors (random slopes and intercepts) in each model. For articulation rate measurements, the number of syllables was also treated as a fixed factor of the model. The maximal models were optimized using stepwise backward elimination [25]. We performed Tukey post-hoc tests utilizing the *emmeans* package [26]. Marginal (R^2m) and conditional R^2 (R^2c) estimates were determined using the *MuMIn* package [27]. For the speaker- and register-wise pause frequencies, after evaluating whether the test assumptions were met, repeated measures ANOVA was performed [28].

3. RESULTS

3.1. Articulation rate

Figure 1 shows the results of articulation rate of IPUs. The analysis revealed that the register (ADS or IDS) in interaction with the baby's age affected the articulation rate ($p < 0.05$, $R^2m = 0.14$, $R^2c = 0.28$).

The Tukey post-hoc test showed that IDS tended to be slower than ADS at 6 and 18 months of the infant's age. (6m: $\beta = 0.21$, SE = 0.07, $z = 2.97$, $p = 0.003$, 18m: $\beta = 0.22$, SE = 0.07, $z = 2.97$, $p = 0.003$). However, immediately after childbirth the two registers did not exhibit articulation rate differences. With increasing syllable number of the IPU, articulation rate also increased ($\beta = 0.09$, SE = 0.03, $t = 27.3$, $p < 0.001$), but the model did not show an interaction between the syllable count and the babies' age or the register. Thus, the articulation rate of IPUs was lower when talking to 6 and 18 month-old infants compared to ADS, even after adjusting to the differences between the syllable counts. Repeating the same analysis on clauses instead of IPUs yielded the same results.

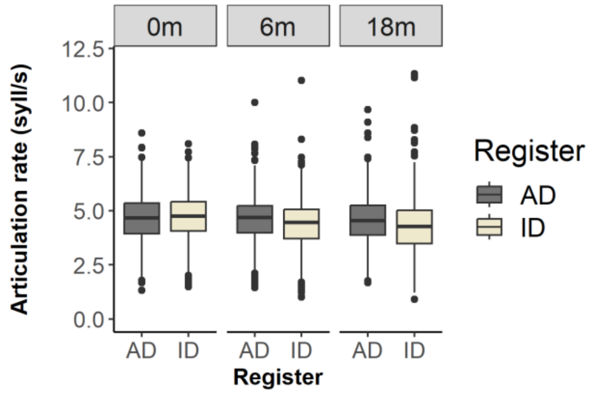

Figure 1: The articulation rate of IPUs for registers and the babies' age.

3.2. Number of syllables

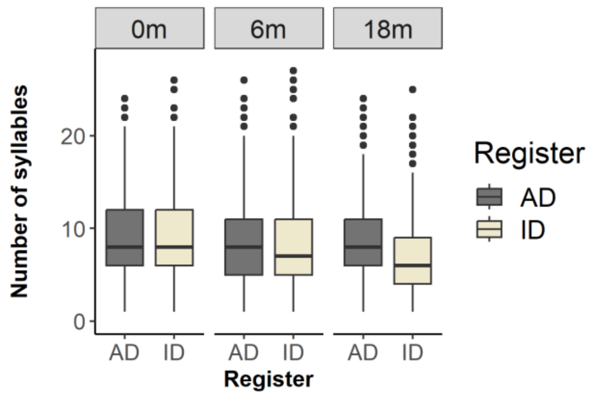

Figure 2: The number of syllables per clauses for registers and the babies' age.

The number of syllables in IPUs and clauses showed a peculiar relationship with the register and the baby's age. The Tukey post-hoc test indicated that ADS exhibited more syllables than IDS both for IPUs and clauses in the experiments conducted 18 months after childbirth (IPU: $\beta = 0.08$, SE = 0.04, $z = 2.20$, $p = 0.028$; clause: $\beta = 0.17$, SE = 0.03, $z = 5.96$, $p < 0.001$) but no such effect was found in the other two investigated ages of the babies. (IPU: $\beta = 0.08$, SE = 0.04, $z = 2.20$, $p = 0.028$; clause: $\beta = 0.17$, SE = 0.03, $z = 5.96$, $p < 0.001$). Although for the syllable counts in IPUs, the model's explanatory power was weak (R^2m = 0.01, R^2c = 0.07), it was found to be moderate for the case of clauses (R^2m = 0.17, R^2c = 0.19), which may well be the consequence of the fact that here the presence or absence of pauses could also be incorporated in the model as a further independent variable.

3.3. Pause duration

Besides register and the babies' age, the position of pauses – whether they occur within or between clauses – constituted a further fixed factor for modeling the effects that influence their durations. The model revealed an interaction between the babies' age and the pause position ($p < 0.05$, R^2m = 0.22, R^2c = 0.24). The duration of pauses was significantly longer between clauses than within clauses for all studied ages of the babies in both registers. The analysis also showed that pauses within the clauses were longer in the 0-month experiments than when the babies were 6 and 18 months old (0m-6m: $\beta = 0.27$, SE = 0.07, $z = 4.19$, $p < 0.001$; 0m-18m: $\beta = 0.36$, SE = 0.07, $z = 5.14$, $p < 0.001$). We found no effect of the register (not even in interaction with the other factors) on the durations of pauses.

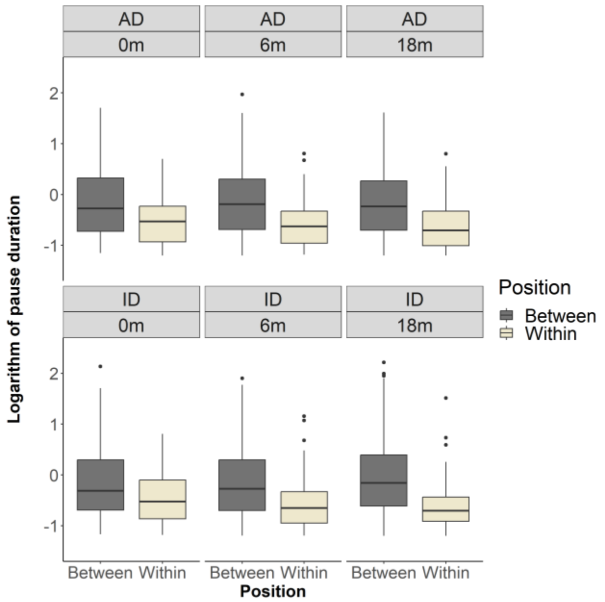

Figure 3: Pause durations for registers and the babies' age between and within clauses.

3.4. Frequency of pauses

The occurrence of pauses (pauses per 100 syllables) was evaluated for each speaker concerning the register and the babies' age. The results indicate a significant interaction between the register and the babies' age (RM ANOVA, $F(1, 24) = 5.51$, $p = 0.01$, $\eta^2 = 0.32$). Tukey post-hoc tests revealed that pauses occurred more often in IDS than in ADS, considering the experiments at 6 and 18 months of the babies' age ($p < 0.05$). However, no such difference between the registers appeared in the 0-month recordings. Post-hoc tests also showed that the pause frequencies in ADS remained unchanged throughout the three recording sessions. To compare pause frequencies between and within the clauses, we calculated the ratio (percentage) of pauses falling between clauses compared to the total number of pauses for each speaker, register, and age. We found that this ratio was higher in IDS than in ADS ($F(1, 12) = 14.99$, $p = 0.002$, $\eta^2 = 0.56$). However, there was no effect of babies' age or interaction with babies' age.

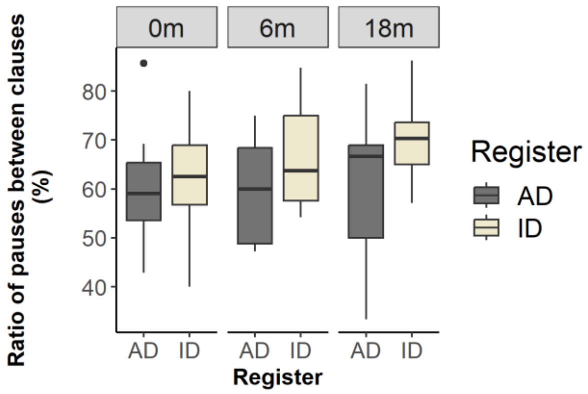

Figure 4: The ratio (percentage) of pauses falling between clauses compared to the total number of pauses for registers and the babies' age.

4. DISCUSSION AND CONCLUSION

In this study, we analyzed the pausing and timing characteristics of infant-directed speech (IDS) and explored the connections between pausing and clause boundaries. Evaluating our records of (semi-) spontaneous narrative IDS and ADS, we found – in agreement with previous research (e.g. [1, 5, 6, 7]) – that clauses and inter-pausal units (IPUs) were both slower in IDS than in ADS, but only in the recording sessions at 6 and 18 months of the babies' age. The observation that this difference did not appear in the 0-month experiments – conducted only a few days after childbirth – could possibly be explained by the special environmental (hospital) conditions and by the lack of actual mother-child interaction in the cases where the infant was asleep during the session [9].

The analyses of the syllable counts in clauses and IPUs revealed that we could observe a difference from ADS only in the sessions with 18-month-old infant addressees in terms of the number of syllables per speech unit. Interestingly, for Mandarin Chinese IDS, it has been reported that speakers tended to use IPUs consisting of less syllables already at two months of the babies' age [11]. The difference may be related to the experimental conditions: their recording sessions were conducted in a way that the babies and their parents were playing freely, whereas in our case, the speech of the subjects was controlled by the arrangement of the experiment (storytelling based on pictures). As for the latter speech situation, it appears that in the earlier ages of the babies, speakers seldom used pauses to isolate and emphasize words and phrases [cf. 29]. Future research involving pragmatic analysis – beyond the scope of the present study – is needed to explore the properties of the shorter IPUs and clauses in the IDS addressing 18-month-olds, who are already capable of producing words from their rapidly increasing vocabulary.

The durations of the pauses did not exhibit differences between the two registers, seemingly contradictory results from earlier research [1, 5, 10]. However, another analysis in which the content was controlled by the experiment setting also did not find any difference between German IDS and ADS in terms of pause durations [12]. Thus, those results and the ones presented in the present paper imply that the speech situation and certain resulting speech planning procedures may have an effect on the characteristics of IDS, including the durations of pauses. However, the duration of pauses was found to be longer between clauses than within clauses for all studied ages of the babies in both registers. Thus, though the durations of pauses were not exaggerated at the clause boundaries in IDS, their lengths still served as boundary markers, and could facilitate the infants' segmentation from fluent speech.

Pauses were more frequent in IDS as compared to ADS at 6 and 18 months of the infants' ages, in agreement with earlier results [7, 12]. Furthermore, the ratio between the frequency of pauses at clause boundaries and within clauses was found to be higher in IDS than in ADS (in agreement with previous research on pauses following utterances in other languages [1, 16]). Therefore, the occurrence of pauses coincided with clause boundaries more often in IDS than in ADS in all three investigated ages of the babies, since pauses may provide crucial cues for the infants to perceive clause boundaries facilitating the acquisition of linguistic structures [3]. It is to be emphasized, however, that although here we have focused merely on a single acoustic cue – pauses – detecting phrase-final lengthening and f0 movements may also help the babies parse the speech. Furthermore, extending the analysis to intonational phrases and their boundaries may provide an even more comprehensive picture of the relations between syntax and prosody in IDS.

5. ACKNOWLEDGEMENTS

We would like to thank Katalin Pirsel and Veronika Harmati-Pap for the manual transcriptions and annotations. This work was funded by the National Research, Development and Innovation Fund (Grants no. NKFI-115385, NKFI-135038 and NKFI-134775).

5. REFERENCES

[1] Fernald, A. Simon, T. 1984. Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology* 20(1), 104.

[2] 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.

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

[4] Broen, P. A. 1972. *The Verbal Environment of the Language-Learning Child*. American Speech and Hearing Association.

[5] Fernald, A., Taeschner, T., Dunn, J., Papousek, Mechthild, de Boysson-Bardies, B., Fukui, I. 1989. A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language* 16(3), 477–501.

[6] Narayan, C. R. McDermott, L. C. 2016. Speech rate and pitch characteristics of infant-directed speech: Longitudinal and cross-linguistic observations. The *Journal of the Acoustical Society of America* 139(3), 1272–1281.

[7] Martin, A., Igarashi, Y., Jincho, N., Mazuka, R. 2016. Utterances in infant-directed speech are shorter, not slower. *Cognition* 156, 52–59.

[8] Han, M., de Jong, N.H., Kager, R.. 2018. Infant-directed speech is not always slower: Cross-linguistic evidence from Dutch and Mandarin Chinese. *Proc. of the 42nd annual Boston University Conference on Language Development*, Somerville. 331–344.

[9] Mády, K., Reichel, U., Szalontai Á., Kohári, A., Deme, A. 2018. Prosodic characteristics of infant-directed speech as a function of maternal parity. *Proc. of the 9th International Conference on Speech Prosody*, Poznań, 294–298.

[10] Stern, D. N., Spieker, S., Barnett, R. K., MacKain, K. 1983. The prosody of maternal speech: Infant age and context related changes. *Journal of Child Language* 10(1), 1–15.

[11] Grieser, D. L. Kuhl, P. K. 1988. Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. Developmental Psychology 24(1), 14–20.

[12] Menn, K. H., Michel, C., Meyer, L., Hoehl, S., Männel, C. 2022. Natural infant-directed speech facilitates neural tracking of prosody. *NeuroImage* 251, 118991.

[13] Soderstrom, M., Blossom, M., Foygel, R., Morgan, J. L. 2008. Acoustical cues and grammatical units in speech to two preverbal infants. Journal of Child *Language* 35(4), 869–902.

[14] Kondaurova, M. V., Bergeson, T. R. 2011. The effects of age and infant hearing status on maternal use of prosodic cues for clause boundaries in speech. *Journal of Speech, Language, and Hearing Research* 54, 740–754.

[15] Ludusan, B., Cristia, A., Martin, A., Mazuka, R., Dupoux, E. 2016. Learnability of prosodic boundaries: Is infant-directed speech easier? *The Journal of the Acoustical Society of America* 140(2), 1239–1250.

[16] Fisher, C. Tokura, H. 1996. Acoustic cues to grammatical structure in infant-directed speech: Cross-linguistic evidence. *Child Development* 67(6), 3192–3218.

[17] Harmati-Pap, V., Vadász, N., Kas, B., Tóth, I. 2021. Anyai dajkanyelvi narratívák lexikai és szintaktikai jellemzőinek longitudinális vizsgálata. *Beszédtudomány - Speech Science* 2, 207–242.

[18] Boersma, P. Weenink, D. 2019. Praat: doing phonetics by computer. Computer program.

[19] Trouvain, J., Werner, R. 2022. A phonetic view on annotating speech pauses and pause-internal phonetic particles. In: Schwarze, C. Grawunder, S. (eds.) *Transkription und Annotation gesprochener Sprache und multimodaler Interaktion: Konzepte, Probleme, Lösungen*., 55–73.

[20] Björkenstam, K. N., Wirén, M., Eklund, R.. 2013. Disfluency in child-directed speech. *Proc. of Fonetik 2013, the XXVIth Swedish Phonetics Conference*, Linköping, 57–60.

[21] Soderstrom, M., Morgan, J. L. 2005. Disfluency in speech input to infants? the interaction of mother and child to create error-free speech input for language acquisition. *Proc. of the DiSS-2005*, Aix -en-Provence, 157–162.

[22] Newport, E. L., Gleitman, H., Gleitman, L. R. 2020. Mother, I'd rather do it myself. Lidz, L. (ed.), *Sentence First, Arguments Afterward: Essays in Language and Learning*. Oxford University Press, 141–177.

[23] R Core Team 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/.

[24] Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. 2017. lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software* 82, 1–26.

[25] Winter, B. 2019. *Statistics for linguists: An introduction using R*. Routledge.

[26] Lenth, R., Singmann, H., Love, J., Buerkner, P., Herve, M. 2022. Package 'emmeans'. Estimated marginal means, aka least-squares means. https://cran.r-project.org/web/packages/emmeans/.

[27] Bartoń, K. 2022. MuMIn: Multi-Model Inference. https://cran.r-project.org/package=MuMIn.

[28] Singmann, H., Bolker, B., Westfall, J., Aust, F., Ben-Shachar, M., S. 2022. Package 'afex'. https://cran.r-project.org/web/packages/afex/

[29] Laing, C. E., Vihman, M., Keren-Portnoy, T. 2017. How salient are onomatopoeia in the early input? A prosodic analysis of infant-directed speech. *Journal of Child Language* 44(5), 1117–1139.