

A Corpus Study of *only* in Child-directed and Child-produced Speech

Lindsay Hracs, University of Calgary

Introduction: Analyzing the distinctive qualities of child-directed speech (CDS) has been a focus of child language acquisition researchers for decades. However, fewer studies (Englund & Behne, 2006; Huttenlocher et al., 2010; Liu et al., 2009; Ramírez-Esparza et al., 2017; among others) have investigated how CDS changes over the course of development. Using corpus analysis, this study seeks to answer the following question: does the frequency of occurrence of *only* in CDS significantly predict the frequency of occurrence of *only* in child-produced speech (CPS)?

Background: *Only* is a canonical example of a Focus Particle (FP). FPs are semantic operators that take scope over a specific constituent in a sentence (König, 1991; Sudhoff, 2010; Beck, 2016; Grosz, 2016), resulting in that constituent being construed as the Focus (i.e. the information a speaker wants a hearer to attend to (Erteschik-Shir, 1973, 1986, 1997)). Consider example (1).

- (1) a. **Only** [Dale]_{FOCUS} drinks coffee. c. Dale drinks **only** [coffee]_{FOCUS}.
b. Dale **only** [drinks coffee]_{FOCUS}. d. Dale drinks [coffee]_{FOCUS} **only**.

The different syntactic positions of *only* result in (i) different c-command relationships; (ii) different intonation patterns; and (iii) different interpretations. Furthermore, learners must come to know that *only* introduces a restriction on a set of possibilities. Thus, context also plays a role.

Previous research on the acquisition of English *only* (Crain et al., 1992; Crain et al., 1994; Kim, 2011; Notley et al., 2009; Paterson et al., 2003; Paterson et al., 2006; Philip & Lynch, 2000) shows that children up to 10-years-old are not adult-like in their usage. However, these studies do not explicitly discuss the role of *input*, i.e. the linguistic stimuli that is analyzed by learners based on the current state their grammars (Carroll, 2017). Instead, empirical claims hinge on participant performance on the experimental stimuli. This study fills a gap in the existing literature by investigating the relationship between CDS and CPS over the course of development.

Methodology: A data set was constructed from the North American English corpora available from the CHILDES database (MacWhinney, 2000). The data set only includes files that contain both CDS and CPS to avoid the criticism that if files are not matched for speaker-type, there can be no measurable relationship between CDS and CPS. A total of 3,040 files from 511 different caregiver-child dyads were included, with child ages ranging from 3 to 117 months. Frequency of occurrence of *only* was extracted for both caregiver and child, for each file, giving 3,040 data points each for CDS and CPS. Since each file had a different total word count, relative frequencies (normalized to 1,000) were used in the analysis.

Results: Overall frequencies are as follows: observed frequency in CDS = 1,788 tokens and CPS = 920 tokens; relative frequency in CDS = 0.409 and CPS = 0.400. Due to the unbalanced and correlated nature of the data, Generalized Estimating Equations (GEE) were used to determine if the frequency of occurrence of *only* in CDS significantly predicts the frequency of occurrence of *only* in CPS. Results show that the occurrence of *only* in CDS is a significant predictor of the occurrence of *only* in CPS, $\chi^2(1) = 11.9, p < .001$. When broken down by age (in 12 month increments), the frequency of occurrence of *only* in CDS significantly predicts the frequency of occurrence of *only* in CPS at 25-36 mos. ($\chi^2(1) = 5.21, p = .022$), 37-48 mos. ($\chi^2(1) = 7.03, p = .008$), 49-60 mos. ($\chi^2(1) = 4.07, p = .044$), 61-72 mos. ($\chi^2(1) = 5.85, p = .016$) and 73-84 mos. ($\chi^2(1) = 36.2, p < .001$).

Conclusion: Although causal claims cannot be made with this corpus study, the results show that in the case of *only*, frequency of occurrence in CDS significantly predicts the frequency of occurrence in CPS. When considering the results by age group, future research must ask what has changed in the children's grammar such that frequency of occurrence in CDS has a significant effect on CPS.

References:

- Beck, S. (2016). Focus sensitive operators. In C. Féry & S. Ishihara (Eds.), *The Oxford handbook of information structure*, 227-250. Oxford: Oxford University Press.
- Carroll, S. E. (2017). Exposure and input in bi-lingual development. *Bilingualism: Language and Cognition*, 20(1), 3-16.
- Crain, S., Drozd, K., Matsuoka, K., Philip, W., & Roeper, T. (1992). *Only in child language*. Paper presented at the 17th Annual Boston University Conference on Language Development.
- Crain, S., Ni, W., & Conway, L. (1994). Learning, parsing and modularity. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 443-467). Hillsdale, NJ : Lawrence Erlbaum.
- Englund, K., & Behne, D. (2006). Changes in infant directed speech in the first six months. *Infant and Child Development*, 15(6), 139-160.
- Erteschik-Shir, N. (1973). *On the nature of island constraints* (Unpublished doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA.
- Erteschik-Shir, N. (1986). Wh-questions and focus. *Linguistics and Philosophy*, 9(2), 117-149.
- Erteschik-Shir, N. (1997). *The dynamics of focus structure*. Cambridge: Cambridge University Press.
- Grosz, P. G. (2016). Information structure and discourse particles. In C. Féry & S. Ishihara (Eds.), *The Oxford handbook of information structure*, 227-250. Oxford: Oxford University Press.
- Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, 61(4), 343-365.
- Kim, S. (2011). *Focus particles at the syntactic, semantic, and pragmatic interfaces: The acquisition of only and even in English* (Unpublished doctoral dissertation). University of Hawaii at Manoa, Honolulu, HI.
- König, E. (1991). *The meaning of focus particles: A comparative perspective*. London: Routledge.
- Liu, H-M., Tsao, F-M., & Kuhl, P. K. (2009). Age-related changes in acoustic modifications of Mandarin maternal speech to preverbal infants and five-year-old children: A longitudinal study. *Journal of Child Language*, 36(4), 909-922.
- MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk (3rd ed.)*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Notley, A., Zhou, P., Crain, S., & Thornton, R. (2009). Children's interpretation of focus expressions in English and Mandarin. *Language Acquisition*, 16(4), 240-288.
- Paterson, K. B., Liversedge, S. P., Rowland, C., & Filik, R. (2003). Children's comprehension of sentences with focus particles. *Cognition*, 89(3), 263-294.
- Paterson, K. B., Liversedge, S. P., White, D., Filik, R., & Jaz, K. (2006). Children's interpretation of ambiguous focus in sentences with "only". *Language Acquisition*, 13(3), 253-284.
- Philip, W., & Lynch, E. (2000). Felicity, relevance, and acquisition of the grammar of every and only. In S. C. Howell, S. A. Fish, & T. Keith-Lucas (Eds.), *BUCLD 24: Proceedings of the 24th Annual Boston University Conference on Language Development* (pp. 583-596). Somerville, MA: Cascadilla Press.
- Ramírez-Esparza, N., García-Sierra, A., Kuhl, P. K. (2017). Look who's talking NOW! Parentese speech, social context, and language development across time. *Frontiers in Psychology*, 8, 1-12.
- Sudhoff, S. (2010). *Focus particles in German: Syntax, prosody, and information structure*. Amsterdam: John Benjamins.

(Note: a list of corpora and/or individual files included in the data set are available upon request.)