

Nonlinear Development of Speaking Rate in Child-Directed Speech Eon-Suk Ko (University at Buffalo, State University of New York)

The quality of speech addressed to young children (CDS; child-directed speech) may play an important role in child language development (Fernald 1985, Thiessen et al. 2005, Hurtado et al. 2008). Despite much research, however, the nature of CDS in relation to the linguistic advances on the part of the child interlocutor is not clearly understood. The current study investigated the developmental trajectory of CDS speaking rate in individual mother-child interactions to help facilitate our understanding of the role of input in child language development. The analysis draws on large scale corpora of both CDS and child speech in the CHILDES database (MacWhinney 2000).

The following corpora were chosen for analysis from the CHILDES database: Thomas (Lieven et al. 2009), Providence (Demuth et al. 2006), Brent (Brent & Siskind 2001) and Soderstrom (Soderstrom et al. 2008). The number of mother-child pairs in the corpora ranged from 1 in Thomas to 16 in Brent corpus. Participants in the Thomas corpus speak British English and participants in the remainder of the corpora speak American English. All children were typically developing, except for Ethan in the Providence corpus who was diagnosed with Asperger syndrome at age 5. The corpora used in this study were chosen because the utterances in their transcripts are linked to media files. From a total of 25 mother-child pairs that included 18 pre-verbal children, a total of 945,255 utterances were analyzed; 638,960 from the CDS and 306,295 from the speech of children.

Speaking rate was calculated based on the transcripts rather than the actual acoustic signal. In the transcripts of the CHILDES corpora chosen, each utterance is marked with two time stamps that correspond with the beginning and end of the utterance in the linked media file; the speaking rate is calculated by dividing the total number of words spoken in an utterance by the timed duration of the utterance. The result is a representation of the speaking rate as number of words per second. To supplement the current method of calculation, an analysis based on semi-automated calculation of speaking rates from the utterances in three one-hour sessions in the Naima data was conducted based on manually segmented clean speech files. The total number of utterances used in this supplementary analysis was 1213.

The main interest in the statistical analysis was to detect any structural changes in the linear relationship between child age and the CDS speaking rate. The current study adopted regression-modeling of the least-square method which results in estimates that approximate the conditional mean of the response variable. In particular, the data were plotted with the breakpoint regression model (Baayen 2008) using R (R Development Core Team 2010), which detects the turning points where the slope of the linear function changes. The most likely breakpoint is determined by first fitting a series of models, one for each possible breakpoint, and then calculating the deviance, the sum of the squared differences between the observed and the fitted values. Finally, the breakpoint for which the deviance is smallest is selected. This model allows for an efficient description of a situation in which there is an abrupt shift in an otherwise linear function. The focus of the analysis was thus to establish whether there is any significant change in the slope, and if so, where this discontinuity is located. ANOVA statistics were used as a measure of testing the null hypothesis of no structural change.

The first analysis was of CDS speaking rates in the corpora of children at an early stage of speech production (Thomas, Naima, Lily, Violet, Alex, William, and Ethan). Raw average

speaking rates of the utterances in each recording session were plotted as a function of age (Figure 1). The number of utterances represented by each of the data points range from 405 to 875.

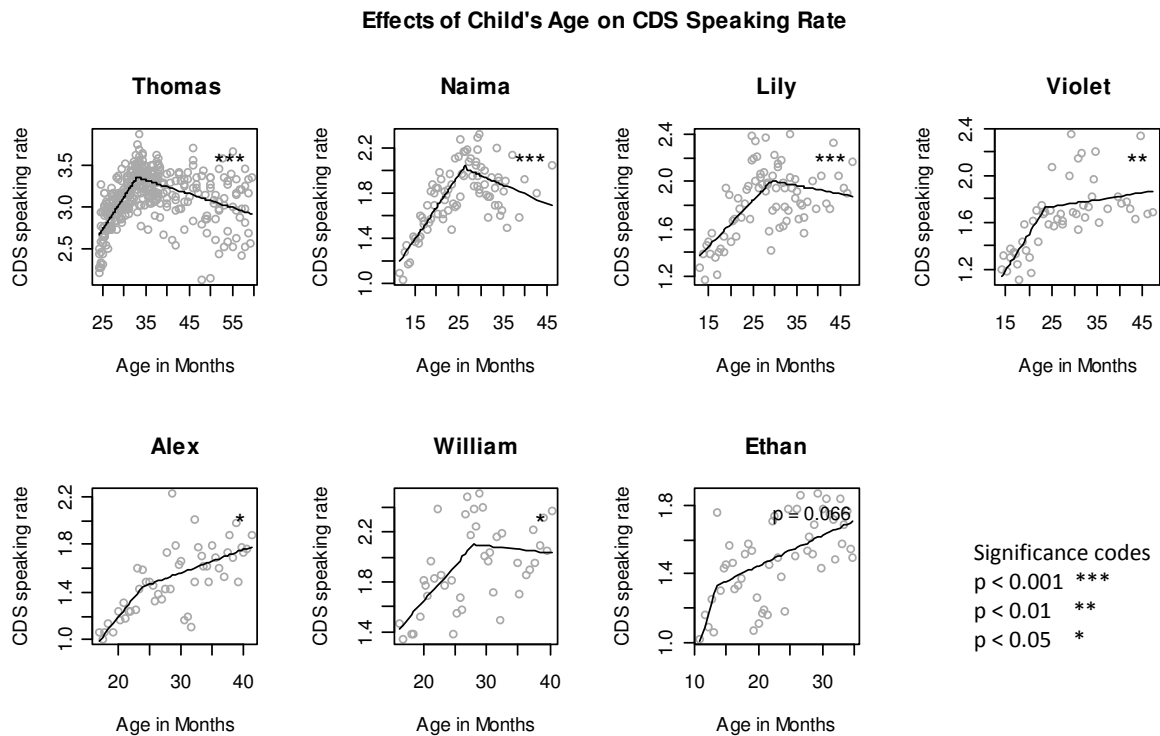


Figure 1. Breakpoint analysis of the CDS speaking rate of mothers as a function of child age. The significance codes are from an ANOVA test comparing the breakpoint model with a simple linear regression model.

What is remarkable from Figure 1 is that the change in CDS speaking rate is nonlinear. ANOVA tests on the breakpoint model by comparing it with a model with just a linear regression line show that the extra parameter for modeling the breakpoint is justified in most mothers' speech analyzed in this study (Figure 1). The null hypothesis of no structural change was rejected in the CDS speaking rate of Thomas ($p < 0.001$), Naima ($p < 0.001$), Lily ($p < 0.001$), Violet ($p < 0.01$), Alex ($p < 0.05$), and William ($p < 0.05$). The significance of the results was marginal in Ethan ($p = 0.066$), whose data may not be directly comparable with that of the other children due to the Asperger syndrome that he was diagnosed with at age 5. The p -values of the ANOVA tests need to be taken with caution because the longitudinal data violate the assumption of independence among observations. Nevertheless, the curvature was so distinct that the existence of discontinuity in the developmental pattern of the CDS speaking rate was evident in most corpora.

The developmental patterns in the speaking rate of the child were analyzed to examine if the CDS speaking rate is mirrored in the speech of the child, and vice versa. In addition, a parallel analysis was made on the MLU of both the mother and child. The results suggest that the nonlinearity of the developmental pattern is present not only in the CDS speaking rate, but also in the speaking rate of the child, and in the MLU of both mother and child (cf. Brown 1973, Scarborough et al. 1986).

Do mothers begin speaking slowly while the child is still pre-verbal or do they reset to the slowest speaking rate with the emergence of speech production in the child? Unfortunately, there is no dense corpus available for duration extensive enough to directly address this question. However, CHILDES database contains data from 18 mothers interacting with their preverbal children, which were analyzed to detect any patterns in the CDS speaking rate in pre-verbal stage. If the acceleration of CDS speaking rate starts during the preverbal stage, we expected to find either evidence for positive correlation between the child's age and mother's speaking rate, or a breakpoint where the slope of the regression line changes from negative to positive. The results suggest that out of the 18 mothers analyzed, only 8 had significant correlation between the child's age and the mother's speaking rate. Out of these 8 mothers' data, 4 of them had negative, and the other 4 had positive coefficient values. The results thus does not provide evidence in support of the hypothesis that the CDS starts out with a low speaking rate and continuously accelerates through the pre-verbal period. A parallel analysis with the mother's MLU yielded similar results.

The alternative hypothesis that mothers' speaking rate is re-set at its lowest value at some point, possibly around the time the child begins to produce his/her own speech, was hard to evaluate given the sparse data. In an independent study, also based on a dense corpus, Roy et al. (2009) found that the MLU of caregivers systematically decrease until about the time the child begins combining words around 16 months, followed by an increase in their MLU's. In the absence of comparable data for the speaking rate, the average speaking rate in the entire data from the 18 mothers in Brent and Soderstrom Corpora was compared with that of the data from the 6 mothers in Providence corpora. The results indicate that the CDS speaking rate during the pre-verbal period (2.89 words/sec) was higher than that in the early speech production period before the breakpoint (1.46 words/sec), $t(297.3) = 24.88, p < 0.001$. This suggests that the possibility of another breakpoint around the time the child begins to produce his/her first words is worth further investigation in future research.

As mentioned earlier, the calculation of speaking rate in this study is based on the number of words in transcripts and the corresponding time stamps rather than the actual acoustic signal. Therefore, the validity of the results based on such a method may need some confirmation. The replication of the nonlinearity in multiple variables in many mother-child pairs suggests that the nonlinearity found in CDS speaking rate is not an artifact of the methodology. The nonlinearity in the MLU is a particularly solid result because of its calculation based on firm word counts, thus the parallel patterns shared between the CDS speaking rate and the MLU corroborates the validity of the findings in this study. Nevertheless, the findings made under current methodology will benefit from a supplementary analysis based on the acoustic signal. Thus, a semi-automated calculation of speaking rates based on speech sound was conducted on the utterances from three critical periods of a randomly chosen mother-child pair, i.e. Naima data.

The utterances were selected at the beginning (age 0;11.28), breakpoint (age 2;1.10), and end (age 3;10.10) of the age range contained in the Naima data. Each session lasted for 81, 82 and 82 minutes, and contained 887, 833, and 643 utterances, respectively. Out of these, 343, 632, and 238 utterances were manually selected after removing utterances that are not appropriate for automated calculation of speaking rate due to the inclusion of overlapping speech, vocal play, background and static noise, etc. The boundaries of these sentences were manually annotated. The speaking rate was then calculated using a Praat script that automatically detects syllable nuclei based on intensity and voicing of the acoustic signal (De Jong & Wempe 2009).

A one-way ANOVA was used to test the differences in mean speaking rates at the three temporal points of Naima data. The mean speaking rates significantly differed across the three temporal points ($F[2, 1210] = 16.62, p < 0.001$). Tukey post-hoc comparisons of the three groups indicate that the speaking rate at age 0;11.28 was significantly lower ($M = 3.13$ syllables/sec) than that at age 2;1.10 ($M = 3.52$ syllables/sec), $p < 0.001$, or that at age 3;10.10 ($M = 3.63$ syllables/sec), $p < 0.001$. However, the difference between the speaking rates at ages 2;1.10 and 3;10.10 was not significant. This result is in line with the findings made earlier in this paper that the CDS speaking rate rapidly accelerates in the beginning of child's early speech production period, but there is a turning point after which it shifts to a different phase.

CDS is often discussed as if there are certain static features that characterize it as a special register, but the findings of this study suggest that the hallmark of CDS is not in its clarity of articulation or slow tempo, as much as in its dynamic changes with shifts occurring around the onset of child speech production and again roughly around 24 months. A parallel pattern of nonlinearity was also observed in the speaking rate of the child and the MLU of both mother and child. These findings suggest that any studies of language acquisition should pay more attention to the dynamic nature of CDS and of child-produced speech, since the nature of the data can be very different depending on whether it is from the pre-verbal, early speech, or

References

- Baayen, R. Harald. (2008) *Analyzing Linguistic Data: A practical introduction to statistics using R*, Cambridge: Cambridge University Press.
- Brent, M. R. & Siskind, J. M (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, **81**, 31-44.
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Demuth, K., Culbertson, J. & Alter, J. 2006. Word-minimality, epenthesis, and coda licensing in the acquisition of English. *Language & Speech*, **49**, 137-174.
- De Jong, N.H. & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior research methods*, 41 (2), 385 - 390.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, **8**, 181-195.
- Hurtado, N., Marchman, V. A. & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science* **11**, F31-F39.
- Lieven, E., Salomo, D. & Tomasello, M. (2009). Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics*, **20**, 481-508.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (third ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Roy, B. C., Frank, M. C. & Roy, D. (2009). Exploring word learning in a high-density longitudinal corpus. *Proceedings of the 31st Annual Meeting of the Cognitive Science Society*. Amsterdam, Netherlands.
- Scarborough, H., Wyckoff, J. & Davidson, R. (1986). A reconsideration of the relation between age and mean utterance length. *Journal of Speech and Hearing Research*, **29**, 394-399.
- Soderstrom, M., Blossom, M., Foygel, I., & Morgan, J.L. (2008). Acoustical cues and grammatical units in speech to two preverbal infants. *Journal of Child Language*, **35**, 869-902.
- Thiessen E. D., Hill, E. A. & Saffran, J. R. (2005) Infant-directed speech facilitates word segmentation. *Infancy* **1**, 53-71.