Adaptation in Child Directed Speech:
Evidence from Corpora

Richard Kunert
Cognitive Science Center Amsterdam
University of Amsterdam
The Netherlands
`rikunert@gmail.com`

Raquel Fernández and **Willem Zuidema**
Institute for Logic, Language & Computation
University of Amsterdam
The Netherlands
`{raquel.fernandez|zuidema}@uva.nl`

Abstract

This paper investigates the dynamics of child-directed speech in longitudinal data from the CHILDES corpus. We quantify the complexity of the speech used by each dialogue participant using simple measures. Our results show that there is a strong correlation in all measures between the complexity of the child's and the mother's utterances, indicating that adults adapt their speech at different levels of linguistic processing when interacting with children in dialogue. These correlations remain, albeit weaker, after controling for the effects of the child's age and for local repetitions in the corpus.

1 Introduction

When adults address young children who are not yet fully competent language users, they use a type of speech that differs from the typical language used in dialogue amongst peers. A large number of studies have shown that *child-directed speech* (CDS) exhibits distinct features at all levels of linguistic processing: it is slower in rate, wider in pitch range, and contains fewer dysfluencies; it is syntactically simpler and less ungrammatical, with shorter sentences and scarce complex structures such as subordinate clauses or sentential complements; and it makes use of a more limited vocabulary, which is typically constrained to the child's interests and focus of attention (see Saxton (2010) for an overview).

Although it is by now uncontroversial that the speech directed to young children constitutes a mode of speaking distinguishable from adult-adult talk, the function and properties of CDS are the subject of considerable debate. One of the main points of

disagreement amongst researchers concerns the extent to which CDS is a necessary condition for language acquisition. Some claim that it is not at all required (Pinker, 1994), while others emphasise its key facilitative role in learning (Dominey and Dodane, 2004) or suggest that it is in fact unavoidable (Saxton, 2009). Another open question regarding the nature of CDS concerns its dynamics. It has been observed that CDS is not a static register but rather a dynamic form of speech that changes over time as the child's language develops – a process referred to as *"finetuning"* by Snow (1995). It is far from clear, however, whether the input to the child is grossly adjusted to the child's age and overall level of development or whether the observed changes are in fact the result of fine-grained adaptations to the child's linguistic behaviour during the course of a conversation.

The present study investigates the dynamics of CDS by investigating adaptation between adults and children in longitudinal data from the CHILDES corpus. We quantify the complexity of the speech used by each dialogue participant (DP) using four simple measures that operate at different linguistic levels – phonology, morphology, lexicon, and syntax – and use correlation analyses to investigate the extent to which the child and adult values of these measures are related. Our results show that there is a strong correlation in all measures as well as in a combined measure of general language complexity, indicating that adults adapt their speech at different levels of linguistic processing when interacting with children in dialogue. We then investigate the potential causes of the observed correlations and the

possible mechanisms driving this adaptation, concentrating on the role of age effects and repetitions.

The paper proceeds as follows: In the next section we review some related work on the dynamics of CDS. In Section 3 we describe the corpus and the methodology we use in our analyses. After that, in Section 4, we report our results and discuss the implications of our findings for models of adaptation in CDS. We finally recap in Section 5.

2 The Dynamics of Child Directed Speech

Child directed speech is often described as a special register – a *motherese* which is significantly different from speech to adults. Since the late seventies, however, researchers investigating the role of linguistic input in the process of language acquisition have noticed that mothers and other caregivers talking to children modify their speech as the child's cognitive and communicative skills evolve (Cross, 1977; Snow, 1989). It is thus well known that CDS changes substantially over time until it becomes standard adult-directed speech. However, despite decades of research, the details of this evolution are not yet well understood – see Snow (1995) and Saxton (2010) for short surveys. An explanation of the dynamic nature of CDS often put forward in the literature is that it changes as a result of an adaptation process of the mother to the child. Snow (1989) refers to this process as *finetuning*, defined as the "adjustment of the level of complexity in CDS in relation to the level of complexity of the child's own output and/or comprehension level."

We can distinguish between a *weak* and a *strong* interpretation of this adaptation process. Under a weak interpretation, the mother would choose a level of speech complexity according to her knowledge of the child's linguistic abilities. Adaptation under this view would be a global process driven by the overall level of development of the child. Some attempts to test this weak version came from studies that compared speech directed to children of different ages using cross-sectional analyses (does speech to 2-year olds differ from speech to 5-year olds?). After conducting one of the most influential studies in this direction, Newport et al. (1977) concluded that mothers did not tune their speech to the developing linguistic abilities of their children. However, the use of cross-sectional data was strongly criticised by Snow et al. (1987), who claimed that in order to test whether a process of adaptation is at play at all, longitudinal studies are required instead.

In contrast to the weak version of the adaptation hypothesis, a strong interpretation suggests that the adaptation process takes place at the micro-level of conversational interaction. Under this view, mother and child align in their contingent responses, with the mother reacting to specific and local cues rather than to global characteristics of the child. This strong version of the hypothesis can be seen as appealing to convergence processes which have been postulated for adult-adult dialogue, such as alignment mechanisms driven by priming effects (Pickering and Garrod, 2004) or coordination mechanisms related to feedback (Brennan and Clark, 1996; Clark, 1996).

Testing the plausibility of the strong interpretation does not only require longitudinal data but also attention to local phenomena. Sokolov (1993) conducted one of the earliest studies that took into account locality by investigating patterns of morphosyntactic usage in adjacent child and parental utterances. The results revealed mutual, local adaptation, which Sokolov argued supported the strong version of the finetuning hypothesis. Syntax is possibly the level of linguistic processing where adaptation effects have been most often demonstrated. For instance, Dale and Spivey (2006) explore the temporal organization of syntactic patterns and conclude that "there is a process of coordination taking place in ongoing conversation at the level of syntactic description", while a few recent studies have found evidence for structural priming in children (Huttenlocher et al., 2004; Gerard et al., 2010).

Here we contribute to current research on the role of adaptation in CDS by providing further evidence that allows us to (1) corroborate in a quantitative way the dynamic character of CDS; (2) test the plausibility of the weak vs. the strong interpretation of the adaptation hypothesis; and (3) study the scope of the adaptation process by looking at different levels of language processing using simple measures.

corpus	# files	number of utterances		
		child	mother	other adults
Adam	55	46733	20354	6344
Sarah	139	38089	29481	16752
Eve	20	12119	10446	4359

Table 1: Total # of dialogues (files), child utterances, and child-directed (adult) utterances in the Brown corpus.

3 Data and Methodology

3.1 Corpus

The CHILDES database (MacWhinney, 2000) contains over 100 corpora of transcriptions of face-to-face dialogues between young children and their caretakers. For the study described in this paper, we use the Brown corpus (Brown, 1973), which includes a total of 214 transcribed longitudinal conversations (each corresponding to a corpus file) with three children, Adam, Eve, and Sarah. The three sub-corpora differ substantially from each other. The Adam corpus contains 55 files with transcripts of conversations recorded over a period of 3 years (age 2;3–5;2); the Sarah corpus covers also approximately 3 years (age 2;3–5;1) and includes more dialogues (139 files) with fewer utterances overall; the Eve corpus is smaller, with only 20 files covering 9 months at an earlier age (age 1;6–2;3).

All files in the three corpora include child-mother interactions. Some files also include additional adult interlocutors, who tend to play a less prominent role (produce fewer utterances than child and mother). Table 1 shows an overview of the overall corpus, and (1) an excerpt from the Adam sub-corpus (adam29):

```
(1) CHI : why it got a little tire?
    MOT : because it's a little truck.
    CHI : can't it be a bigger truck?
    MOT : that one can't be a bigger truck
          but there are bigger trucks.
```

3.2 Measures of Speech Complexity

We use four simple measures to quantify the complexity of the speech used by each dialogue participant at different levels of linguistic processing.

- Mean Utterance Length (UL): length of utterance measured in words, averaged over a dialogue;[1] this is a rough indicator of syntactic complexity.

- Mean Word Length (WL): length of words measured in characters, averaged over a dialogue; a rough indicator of morphological complexity.

- Mean Number of Word Types (WT): the number of distinct word types in a dialogue divided by the number of utterances by the relevant speaker in that dialogue; a rough indicator of lexical complexity.

- Mean Number of Consonant Triples (CT): the number of consonant triples (in the surface orthographic form) per utterance per dialogue; a rough indicator of phonological complexity.

All four measures are, admittedly, rather crude approximations of the underlying linguistic complexities, but they are straightforward to use and turn out to suffice for our purposes. Thanks to the large sizes of the corpora we use, they yield clear patterns when tracking their evolution with child age and correlations between, e.g., the child's and mother's utterances (see Section 4).

Additionally, we combine the four measures above to obtain a measure of the overall language complexity. The combined measure can be thought of as a kind of average of the four basic measures; it is obtained by summing after transforming the four measures to the common scale of z-scores, where all values are expressed as distance, in standard deviations, from the mean.

- General Complexity (GC): the sum of UL, WL, WT and CT, after applying the z-score-transform to each.

That is, for a dialogue i and speaker j, general complexity $GC(i,j) = z_{UL}(i,j) + z_{WL}(i,j) + z_{WT}(i,j) + z_{CT}(i,j)$, where $z_X(i,j) = \frac{X(i,j) - \mu_X(j)}{\sigma_X(j)}$, with $X \in \{UL, WL, WT, CT\}$, $\mu(j)$ the mean of values for j over the entire corpus and $\sigma(j)$ the standard deviation of these values.

---

[1]This measure differs from the commonly used Mean Length of Utterance (MLU), corresponding to the mean number of morphemes per utterance. Morphological complexity, in our approach, is measured by mean length of words (WL).
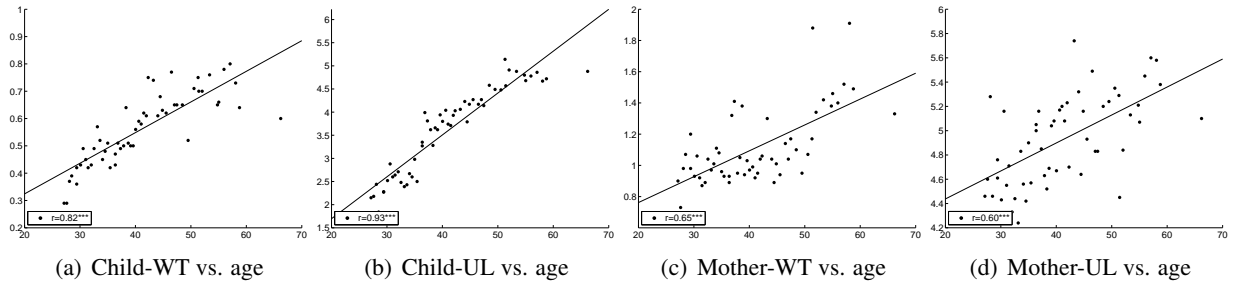
(a) Child-WT vs. age   (b) Child-UL vs. age   (c) Mother-WT vs. age   (d) Mother-UL vs. age

Figure 1: Scatter plots and regression lines, showing the change in complexity of child utterances (a,b) and the mother's child directed speech (c,d) with age of the child (in months). Shown are the measures of syntactic (UL) and lexical (WT) complexity defined in section 3.2.

3.3 Complexity against Age

Having defined our complexity measures, the first question to ask is whether they indeed allow us to track the clear increase in complexity during the child's development. This is indeed the case; in Figure 1(a,b) we illustrate the well-known but nevertheless impressive increases in vocabulary size and utterance length with age, with data obtained from the Adam corpus. Adam's vocabulary becomes much richer, rising from 0.3 new words per utterance (i.e., he uses a word that he hasn't used before in the current dialogue only once every 3 utterances) at age 27 months, to close to 1 new word per utterance at age 57 months. And Adam's sentences go from an average length of 2 words to an average length of about 5 words in the same period.

The second question to ask is whether and how the mother's utterances change in complexity over the same time period. In Figure 1(c,d) we show how the complexity of the mother's utterances changes with the age of the child. As is clear from this figure, the utterances of the mother also undergo a clear – though less radical – development, rising from 1 to about 2 novel words per utterance, and from utterances of length 4.5 to utterances close to 6 words long. Hence, these data illustrate another well-known finding from child language research pointed out in Section 2: that child-directed speech is not a fixed register, insensitive to the abilities of the child, but that it changes as the child develops with age.

Moreover, the plots in Figure 1 point to a third aspect of child and child-directed speech: the developments in the child and the mother appear to be highly, but not perfectly correlated. In this study,

we take this observation as our starting point and investigate the extent to which the complexity of the child's and caretaker's utterances are correlated, and what the possible causes of these correlations are.

3.4 Measuring Correlations

Our five variables (UL, WL, WT, CT, and GC) are global measures of speech complexity computed on a per-dialogue basis: for each dialogue in the corpus and for each interlocutor taking part in that dialogue, we calculate one value for each variable. Our interest is in investigating whether child and adult values of these variables correlate using the Pearson product-moment correlation coefficient (Pearson's r). Hence, for a measure X and a pair of speakers $\langle j, k \rangle$ we calculate:

$$r(X, j, k) = \frac{1}{n-1} \sum_{i=1}^{n} z_X(i, j) \cdot z_X(i, k)$$

Since, as mentioned above, some adult interlocutors play a very minor role in some conversations, in all correlations between pairs of dialogue participants reported in the experiments described in the next section, we consider only those dialogues where each dialogue participant in the relevant pair contributes at least 50 utterance.

4 Analyses & Results

4.1 Baseline results

Figure 2 contains scatter plots based on child and mother utterances from the Adam corpus, showing strong correlations between the values on each of our four basic measures. Figure 3(a) summarizes this data, showing Pearson's r-values for each of
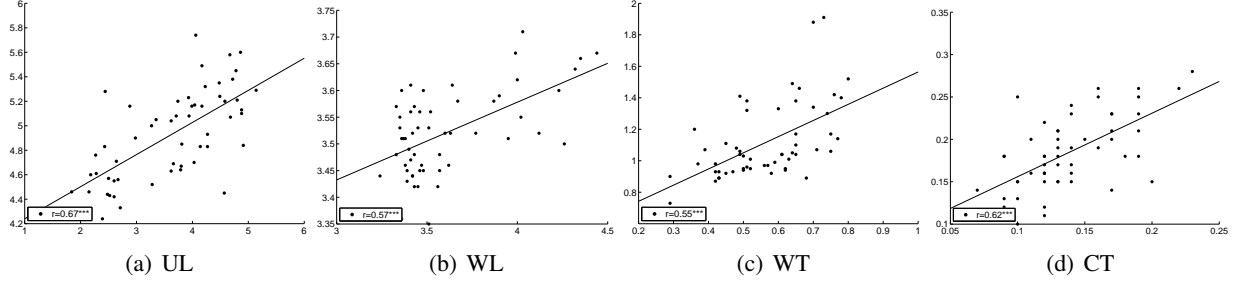
(a) UL (b) WL (c) WT (d) CT

Figure 2: Scatter plots, showing the relation between complexity of child utterances (horizontal axis) and the mother's child directed speech (vertical axis) in the Adam corpus. Each dot thus represents data from 1 file in the corpus. Shown are the measures of syntactic, morphological, lexical and phonological complexity defined in section 3.2.
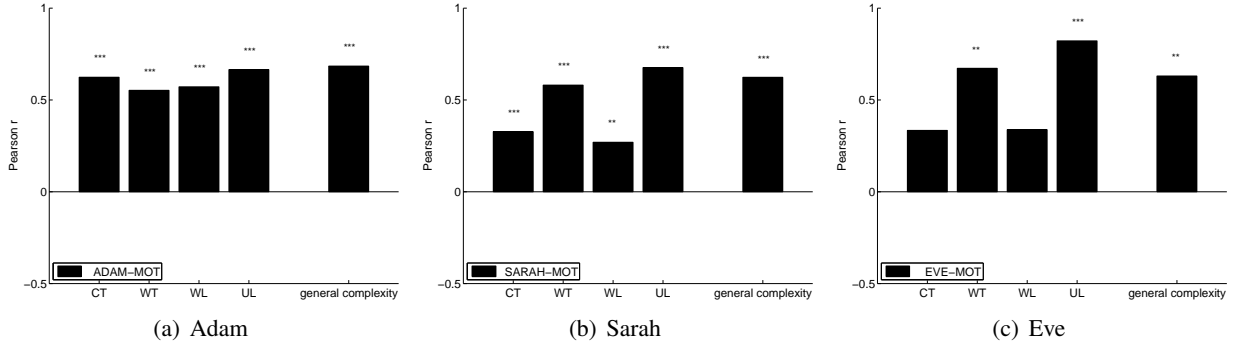

(a) Adam (b) Sarah (c) Eve

Figure 3: Correlations between the child's and mother's utterances on each of the five complexity measures, for each of the three children.

the four pairwise correlations, plus the r for the correlation in general complexity between Adam and his mother. Asterisks indicate that all correlations are highly significant.[2] Figures 3(b,c) give the same data for Sarah and Eve, and show that the correlations are robust across measures and child-mother pairs.

4.2 Controlling for Age

How do we explain the strong correlations we find? An obvious candidate answer is that they emerge from the fact that both the child's complexity and the mother's complexity increase with age. For the child this is a direct consequence of the acquisition process, but for the mother several mechanisms could be proposed that differ in whether or not the interaction with the child *within* the dialogue determines the mother's complexity level. A mechanism that operates without appeal to such dialogue principles would involve the mother choosing the complexity level of her speech based on her knowledge of

the child's developmental stage. Under this model, which would support the weak interpretation of the adaptation hypothesis mentioned in Section 2, we expect no correlations beyond those explained by age. On the other hand, a mechanism that does involve dialogue principles and that is thus in line with a strong interpretation of adaptation would predict that child and mother complexity are correlated over and above the effects of age.

In order to test this we used *partial correlations*. This method removes the common variance shared by the child's values, the mother's values and child age and takes the remaining common variance of the child's and the maternal values as the basis for the correlation coefficient. Hence, the r value of a correlation between a pair of speakers $\langle j, k \rangle$ after controlling for age (A) corresponds to:

$$r_{j,k,A} = \frac{r_{jk} - r_{xA}r_{yA}}{\sqrt{(1 - r_{jA}^2)(1 - r_{kA}^2)}}$$

This technique can be thought of as a way of correlating two variables while holding age constant.

---

[2] The convention we use to indicate significance levels is: *** $p < .001$, ** $p < .01$, * $p < .05$.
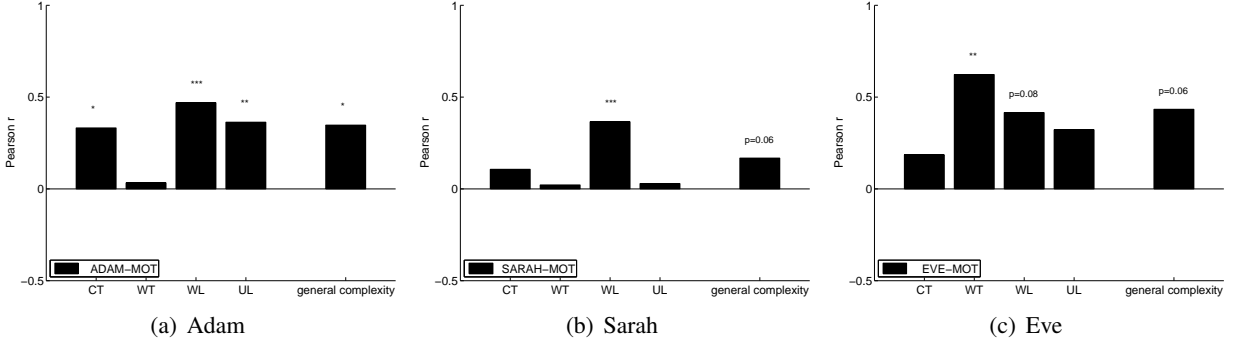
(a) Adam   (b) Sarah   (c) Eve

Figure 4: Partial correlations between the child's and mother's utterances on each of the five complexity measures, for each of the three children, after controlling for the effects of age.

Figure 4 shows that the remaining variability in child values and maternal values is sufficient to still produce significant correlations. This suggests that the baseline findings cannot be entirely accounted for by child age, and that the mother does not only adapt to the general developmental stage of the child (as represented by age) but also to her child's performance in a given interaction.

4.3 Controlling for Repetition

As is well known, one of the features that characterises the interaction between young children and adults is a significant level of repetition. Adults addressing children repeat themselves a lot and they also repeat the child's speech, often with minor variations to the original utterance. Likewise, children repeat utterances they hear from caretakers. If such repetitions lead to a significant number of highly similar utterances of mother and child, then they are also responsible for part of the correlations on the complexity measures that we observed.

Note that the conversational mechanisms that yield repetitions – such as recasts, clarification requests, or priming effects – typically operate on (near-)contingent utterances, and thus form the simplest instance of the strong finetuning hypothesis of child-directed speech. It is important to establish whether this simple explanation suffices to explain the observed correlations, or whether additional, more sophisticated conversational mechanisms need to be assumed.

To investigate these issues, we try to isolate the contribution of repetitions. We cannot, however, control for repetition in the same way that we controlled for age: where age of the child can reason-
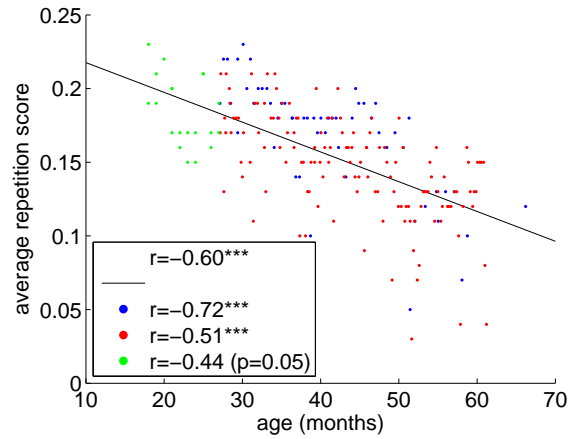

Figure 5: Repetitiveness decreases with age. Shown are average repetiveness scores per dialogue for utterances from Adam (blue), Sarah (red) and Eve (green).

ably be expected to have a causal effect on linguistic complexity, repetition has no direct effect on complexity. Rather, repetition directly affects how correlated the child's and caretaker's complexities are (regardless of whether they are high together or low together).

Our approach consists of two steps. First, we calculate a *repetition score* ρ per utterance, that can be thought of as a way to quantify the likelihood that a particular utterance constituted a repetition. Second, we calculate a threshold value for repetition scores, and remove utterances with scores higher than the threshold from the corpus. The threshold is calculated so as to avoid both false positives and false negatives, as we explain below. We then recalculate the values of our five measures and redo the correlation analyses.

In the procedure to obtain repetitions scores, we calculate the similarity between an utterance u and the preceding utterances v of the other speaker. We then discount this similarity with the distance in the dialogue between u and v, such that the highest values are obtained if u and v are both very close and very similar. The final repetition score $\rho(u)$ of utterance u is the maximum of the discounted similarities with all earlier utterances of the other speaker. Thus:

$$\rho(u) = \max_{v:t(v)<t(u)} s(u,v)c^{d(u,v)}$$
$$s(u,v) = 1 - \frac{L(u,v)}{\max(|u|,|v|)}$$
$$d(u,v) = |\{v'|t(v) < t(v') < t(u)\}|$$

where $t(u)$ refers to the time that u is produced (measured in line numbers in the corpus); $s(u,v)$ is the similarity between u and v, obtained by calculating Levenshtein distance $L(u,v)$, dividing by its maximum value (the length of string u or v) and subtracting the result from 1; $d(u,v)$ gives the distance between u and v in the dialogue, measured in number of utterances v' in between u and v (with v and v' produced by the same speaker); $c = 0.9$ is an arbitrary constant (just below 1).

We computed repetition scores for all utterances by the child and the mother in the corpora of the three different children. To assess whether this way of quantifying repetition is meaningful, we check whether we can reproduce the well-known phenomenon of decrease in repetitiveness with age. The scatter plot in Figure 5 shows the average repetition scores of the child's and mother's utterances per dialogue against child age. The graph clearly shows that the degree of repetitiveness decreases with age, consistently for the three children.

To determine the threshold θ above which an utterance is classified a 'repetition' and removed from further analysis, we randomised the utterances in the entire Adam corpus and calculated the repetition score per utterance of the Adam-mother interaction. In the randomised corpus, true repetitions are very unlikely to occur near to their source; the distribution of repetition scores in the randomized corpus therefore tells us for different thresholds what the likelihood of false positives is.

We choose a repetition threshold at two standard deviations above the mean of the randomised cor-
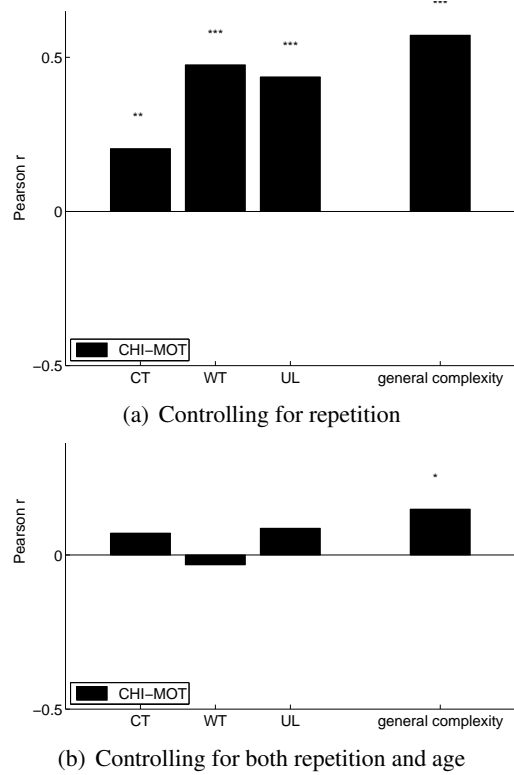

(a) Controlling for repetition

(b) Controlling for both repetition and age

Figure 6: Correlations between all children's and their mother's utterances on all measures except WL.

pus's repetition score distribution (thus allowing for about 2.2% false positives). For the given arbitrary constant $c = 0.9$, this gives a threshold at $\theta = 0.3$. Controlling for repetition in this manner reduced the number of data points substantially: 22% of all child or mother utterances are discarded over the three corpora (Adam: 24%, Sarah: 19%, Eve: 27%).

Note that we calculate repetition scores for entire utterances; this way of controlling for repetition thus makes most sense for our non-word-based measures (UL, WT, and CT). We considered defining a similar control at the word level, where indiviual words are deleted if they are too similar to a word just used by the other speaker. But this approach fails due to the high frequency of function words, and we have so far not found a good alternative. We therefore leave aside WL in the results in this section.

Results for the correlation analysis after controlling for repetition are in Figure 6(a), where we show the results for the three children aggregated. We find that despite the conservative definition of repetitions, the pattern of results is similar to the one observed in the baseline results (Figure 3). While

some variable-correlations no longer reach significance (Adam corpus's WT and UL, Eve corpus's CT), the general complexity score-correlations are significantly positive ($p < .01$) for all three child-mother dyads. Hence, repetition alone does not explain the observed correlations either, even though repetitions are frequent in the data and make the correlations stronger.

How much of the correlations still observed after the repetition control might be due to the effects of the age of the child? Figure 6(b) reports the correlations we find after controlling both for repetition and for age. With both controls, the results are less clear than before. Some correlations almost disappear and some become insignificant. However, when the data is pooled across the three children, the general complexity measure remains significant after controlling for the effects of age.

5 Conclusions and Future Directions

We have investigated the dynamics of CDS by quantifying the complexity of the speech used by each dialogue participant by means of simple measures that operate at different levels of linguistic processing. Our results show that there are strong correlations in linguistic complexity between the child and mother utterances. We have demonstrated that these correlations are only partly explained by the child's age and the local repetitions that characterise child-adult dialogues. These results lend support to the strong version of the finetuning hypothesis on child-directed speech. There remain some significant correlations after controlling for age and repetition using the current coarse-grained method, calling for further investigation of local dialogue mechanisms in mother-child interactions.

Acknowledgements

References

S. Brennan and H. Clark. 1996. Conceptual Pacts and Lexical Choice in Conversation. *Journal of Experimental Psychology*, 22(6):1482–1493.

R. Brown. 1973. *A first language: The early stages.* Harvard U. Press.

H. Clark. 1996. *Using language*. CUP.

T.G. Cross. 1977. Mothers' speech adjustments: The contribution of selected child listener variables. In C. Snow and C. Ferguson, editors, *Talking to children: Language input and acquisition*, pages 151–188. CUP.

R. Dale and M.J. Spivey. 2006. Unraveling the dyad: Using recurrence analysis to explore patterns of syntactic coordination between children and caregivers in conversation. *Language Learning*, 56(3):391–430.

P.F. Dominey and C. Dodane. 2004. Indeterminacy in language acquisition: the role of child directed speech and joint attention. *Journal of Neurolinguistics*, 17(2-3):121–145.

J. Gerard, F. Keller, and T. Palpanas. 2010. Corpus evidence for age effects on priming in child language. In *Proceedings of CogSci*, pages 1559–1564.

J. Huttenlocher, M. Vasilyeva, and P. Shimpi. 2004. Syntactic priming in young children. *Journal of Memory and Language*, 50(2):182–195.

B. MacWhinney. 2000. *The CHILDES Project: Tools for analyzing talk (Vol. 2: The Database)*. Erlbaum.

E.L. Newport, H. Gleitman, and LR Gleitman. 1977. Mother, I'd rather do it myself: Some Effects and noneffects of maternal speech style. In C. Snow and C. Ferguson, editors, *Talking to children*, pages 109–49. CUP.

M.J. Pickering and S. Garrod. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(02):169–190.

S. Pinker. 1994. *The language instinct: The new science of language and mind*. Penguin.

M. Saxton. 2009. The inevitability of child directed speech. In *Advances in language acquisition*, pages 62–86. Palgrave Macmillan.

M. Saxton. 2010. *Child language: Acquisition and development*. Sage Publications Ltd.

C.E. Snow, R. Perlmann, and D. Nathan. 1987. Why routines are different: Toward a multiple-factors model of the relation between input and language acquisition. In *Children's language*, volume 6, pages 65–97.

C.E. Snow. 1989. Understanding social interaction and language acquisition; sentences are not enough. In *Interaction in Human Development*, pages 83–103.

C.E. Snow. 1995. Issues in the study of input: Fine-tuning, universality, individual and developmental differences, and necessary causes. In *The Handbook of Child Language*, pages 180–193. Blackwell.

J.L. Sokolov. 1993. A local contingency analysis of the fine-tuning hypothesis. *Developmental Psychology*, 29(6):1008–1023.