Linguistic input is coordinated to children's developmental level

Daniel Yurovsky
yurovsky@stanford.edu
Department of Psychology
Stanford University

Gabriel Doyle
gdoyle@stanford.edu
Department of Psychology
Stanford University

Michael C. Frank
mcfrank@stanford.edu
Department of Psychology
Stanford University

Abstract

Children rapidly learn a tremendous amount about language despite limitations imposed on them by their developing cognitive processes. One possibility is that caregivers *coordinate* the language they produce to these limitations, titrating the complexity of their speech at developmentally-appropriate levels. We test this proposal by measuring the extent to which parents alter their speech to children in a contingent manner over the course of the first 5 years. Our large-scale corpus analysis confirms this prediction, showing a high degree of mostly parent-led coordination early in development that decreases as children become more proficient language learners and users.

Keywords: Language acquisition, cognitive development, computational models

Introduction

Children learn a tremendous amount about language in their first few years of life. By the time they are able to run down the street, typically developing children have over a thousand words in their productive vocabularies (Mayor & Plunkett, 2011). They can combine these words to produce new, meaningful multi-word utterances (Lieven, Salomo, & Tomasello, 2009). They can even learn new words from just the syntactic constructions in which they occur (Yuan & Fisher, 2009). What explains this rapid developmental progression?

Infants have available to them a surprising range of learning mechanisms. By 8-months, children can use distributional properties of language to segment discrete words from continuous speech (Saffran, Aslin, & Newport, 1996), and by 12 months can use these same kinds of cues to learn ordering regularities in artificial grammars (Gomez & Gerken, 1999) and mappings between words and objects (Smith & Yu, 2008). But while these and other competencies are available early, children's level of learning *performance* is often strikingly limited. For instance, children's learning of new words from distributional properties of language is highly constrained by their developing attentional and memory systems (Vlach & Johnson, 2013).

Why do children learn so quickly if their learning is so constrained? One possibility is that child-directed speech differs systematically from speech used to test their learning in the laboratory. Indeed, the language that parents produce to their children—across a variety of levels and structures—appears to contain many redundant cues and regularities that facilitate learning (Gogate, Bahrick, & Watson, 2000; Thiessen, Hill, & Saffran, 2005; Yurovsky, Yu, & Smith, 2012). However, in some cases child-directed speech appears systematically different in ways that do not support learning. For instance, child-directed speech typically contains simpler and less variable syntactic structures. This simplicity is thought to aid in early grammatical acquisition, but also makes it *harder* to learn more complex constructions (Montag, Jones, & Smith, 2015; Montag & MacDonald, 2015). [mcf: This doesn't make total sense to me - is the point that this input is tailored for younger kids but not older kids, while the other examples are tailored for both?]

The solution, then, may be neither in the learner nor in the input, but in the coordination between learner and input. This hypothesis is intuitively appealing, and some early evidence supported the idea that parents tune the complexity of their language to the developing abilities and needs of their children (Snow, 1972). But two pieces of evidence speak against the strongest form of the *linguistic tuning* hypothesis. First, parents do not appear to use simpler words when speaking to younger children (Hayes & Ahrens, 1988). Second, parents rarely correct their children's syntactic errors, and children are resistant to the few corrections they get (Brown & Hanlon, 1970; Newport, Gleitman, & Gleitman, 1977).

More recent work has suggested other—perhaps more subtle—ways in which parents might tune the language they produce in response to children's competence. Although they do not correct syntactic errors, parents are much more likely to repeat and reformulate children's ungrammatical utterancs, providing a potential corrective signal (Hirsh-Pasek, Treiman, & Schneiderman, 1984; Chouinard & Clark, 2003). And although parents' vocabulary choices may be due at least in part to the content they need to convey, there may still be ways for adults to tailor the structure of their utterances to children's perceived competence. In the current paper we pursue this latter hypothesis.

To test whether parents tune their speech to that of their children, we conduct a large-scale corpus study using *linguistic alignment*, a measure of how much speakers change the way they talk to accommodate their conversational partners. Critically, alignment is a local measure: High alignment results not from choosing words that are simpler overall, but from choosing words that easier to process in context (c.f. Hayes & Ahrens, 1988). We predict that caregivers should align more to their younger children, altering their speech more when children need more linguistic support.

Linguistic alignment

When we use language to communicate, we are trying to use the words we say to convey the message we intend. Some of the words will be obligatory to getting the message across. Consider the conversation in Table 1, taken from the Providence corpus (Demuth, Culbertson, & Alter, 2006). In her first response, Naima's mom has little choice but to say

Naima:	Eating **that**. Eating **some of that**.
Mom:	**Some of this**? **You** know what **that** is? **It** is sweet potato.
Naima:	**I** am **a** bear **that** eats.
Mom:	**You're a** bear **that** eats what? What do **you** eat little bear?
Naima:	Fresh pear

Table 1: A snippet of the transcript between Naima (at 20-mo.) and her mother in the Providence Corpus. Bolded words are members of the LIWC categories and were included in the model (Pennebaker et al., 2007).

"sweet potato" if she wants to inform Naima that they are eating sweet potato. However, she could perfectly well have left out the words "some," "that," and "this," or exchanged them for others and still conveyed the identity of the food on Naima's plate.

Speakers are influenced in their production choices by a tremendous range of factors, ranging from the phonological to the sociolinguistic, but we focus here one particular reason for these choices: contingency on a conversational partner. When we talk, we tend to re-use use each-other's expressions, aligning to each other. This kind of alignment appears to be a pervasive property of human social interaction and linguistic communication (Giles, Coupland, & Coupland, 1991; Garrod & Pickering, 2004). Further, this alignment appears to be useful, facilitating fluent processing of speech, and increasing the probability of successful communication and accomplishment of joint goals (Ireland et al., 2011; Fusaroli et al., 2012). Critically, alignment is directional: even in the same conversation, some speakers will align more than others. For instance, alignment varies across a social hierarchy, with less powerful speakers aligning more to powerful speakers (Kacewicz, Pennebaker, Davis, Jeon, & Graesser, 2013). Thus, linguistic alignment can measure a speaker's effort to coordinate with a conversational partner. We leverage this property to measure the extent to which parents are altering the way the speak to coordinate with their developing children.

In our analysis, we explicitly focus on the words that are least critical for conveying the content of the message. If Mom shows an increased likelihood of using content words in a way that is contingent on the content words produced by Naima, we can conclude only that they are talking about the same thing. However, if Mom is more likely to say "sweet potato" after Naima says it, we can conclude only that they are talking about the same thing. However, if mom is more likely to say "this" after Naima does, she is more likely to be changing her style in a way that makes it easier for Naima to process. [mcf: This bit (two prev sentences) seemed like it could be cut down substantially.] To capture this idea, we chose as our target words a set of 676 words falling into 14 categories that are identified as "strictly non-topical style dimension" (Linguistic Inquiry and Word Count; Pennebaker et al., 2007). We perform our analyses at the level of categories, to capture both exact repetitions of a conversational partner's words and also reformulations and expansions (Chouinard & Clark, 2003). Thus, in Mom's response, both "this" and "that" would count as instances of the impersonal pronoun category (Table 2). These 14 LIWC categories have been used by us and others in previous work examining alignment in a variety of contexts from social media to supreme court proceedings (?, ?, ?).

Model

To estimate linguistic alignment between parents and their children, and how it changes over development, we extended the model introduced by ? (?) to measure linguistic alignment in large, sparse datasets. The goal of this model is to estimate the extent to which a speaker's use of function words is influenced by their conversational partner's use of words in the same category. In Mom's first response to Naima, for instance, she uses the quantifier "some" and the indefinite pronoun "this" (Table 1). The question our model addresses is whether Mom would have been less likely to use these words if Naima had not produced a quantifier and an indefinite pronoun in her previous utterance?

The model attempts to predict on an utterance-by-utterance level whether a speaker will produce a word belonging to each of the 14 function word categories. The probability of producing a word is controlled by two independent factors: (1) the speaker's baseline probability of producing that word, and (2) the speaker's tendency to align to their conversational partner, producing words from categories that their partner just produced. Thus, the primary computation in the model is essentially the same as standard logistic regression, which predicts a binary response (here, production) via the sum of a set of predictors (here, word frequency and alignment probability). The remaining machinery of the model allows frequency and alignment estimates to be pooled hierarchically across different categories, speakers, and ages in a principled way.

Model Details

The full graphical representation of our model is shown in Figure 1. The model operates over a representation of utterances as binary vectors indicating the presence or absence of each of the 14 LIWC categories. The probability of producing each category in each utterance is then predicted via two parameters: (1) The speaker's baseline probability for using that word category (η^{base}), and (2) The speaker's change from this baseline due to interacting with the listener (η^{align}). For messages not following a category's use, the categories parameter is produced by taking the inverse logit of it's baseline log odds $\left(logit^{-1}\left(\eta^{base}\right)\right)$. If the utterance follows an utterance that contains the category, we say that it's probability of production is the inverse logit of the sum of the baseline and alignment log odds $\left(logit^{-1}\left(\eta^{base}+\left(\eta^{align}\right)\right)\right)$.

Because the LIWC categories vary widely in the production frequencies, we draw the log odds of each from an in-

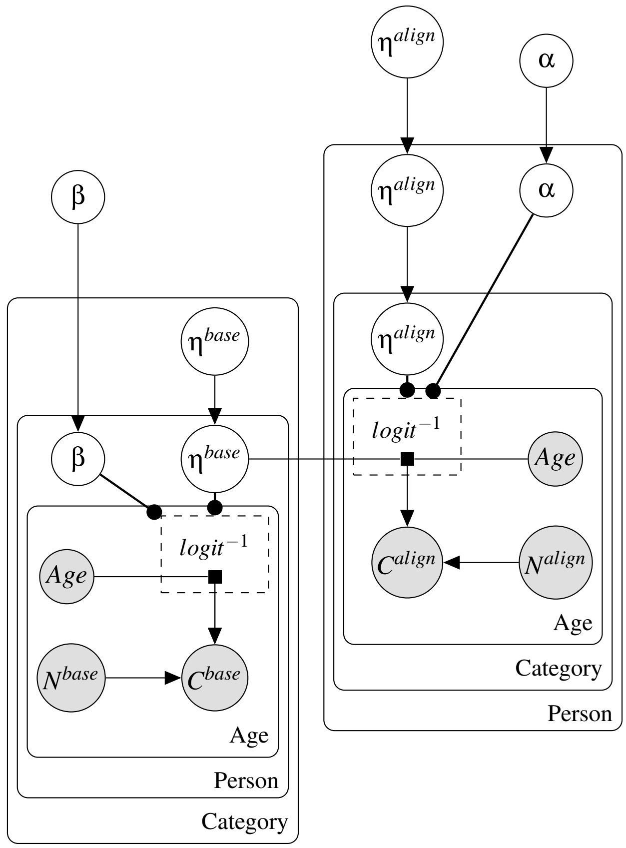

Figure 1: The Hierarchical Alignment Model (HAM) we use to analyze linguistic alignment in CHILDES. Speaker's word choices are modeled as having two influences: (1) Their baseline probability of using each word category (η_{base}), and (2) their increase from baseline due to their conversational partner's use of each category in the previous utterance (η_{align}).

Category	Examples	Adult	Child
Article	*a, an, the*	.31	.13
Certainty	*always, never*	.05	.01
Conjunction	*but, and, though*	.22	.06
Discrepancy	*should, would*	.11	.04
Exclusive	*without, exclude*	.12	..04
Inclusive	*with, include*	.21	.07
Indefinite pronoun	*with, include*	.45	.19
Negation	*not, never*	.21	.11
Preposition	*to, in, by, from*	.38	.14
Quantifier	*few, many*	.13	.04
Tentative	*maybe, perhaps*	.11	.03
1st person singular	*I, me, mine*	.05	.04
1st person plural	*we, us, ours*	.09	.02
2nd person pronoun	*you, yourself*	.33	.06

Table 2: Marker categories for linguistic alignment, with examples and probability of appearing in adult and child productions.

dependent uninformative uniform prior ($Uniform(-5,5)$), which covers more than the range of observed probabilities without putting too probability much mass on extremely large or small values. We put a conservative prior on alignment, drawing $\eta^{align} \sim Normal(0,.5)$, regularizing it strongly towards zero.

To pool data across participants for robust estimation, we estimate all parameters hierarchically. We say that there is a population-level of alignment, which generates speaker-levels of alignment, which generate category-levels of alignment. This decision allows us both to make principled inferences both about how much parents align to their children in general, and about how much specific parents align to their children. For baseline probabilities, we instead nest people within categories as the empirical baseline production probabilities vary widely across categories.

Finally, we let both the probability of using any of these categories (β), and the probability of aligning (α) vary over development. Inspection of posterior parameter estimates showed alignment varied approximately linearly with age, so for simplicity, we formalize β and α as linear scalars of η^{base} and η^{align} respectively.

Using this model, we can test two distinct hypotheses about the way that parents might coordinate to their children's developmental level. First, if β is non-zero, then we can infer that speakers change their baseline likelihood of producing the function words in the 14 LIWC categories over the course of children's development. If parents' β is positive, they are using simpler words earlier in development. Second, if α is non-zero, then we can infer that speakers change their level of alignment over development. If parents' α is negative, they are aligning more to their younger children.

Simulations

Data

To maximize the power and generalizability of our analysis, we selected all of the English-language transcripts available in CHILDES (?, ?) that contained conversations between a parent and a target child who was 12–60-mos. of age. This criterion resulted in a total of 3,851 transcripts across 417 unique children. The number of transcripts per child varied widely, ranging from 1 ($n = 164$) to 440 ($n = 1$), with a median of 26.

For each transcript, we first combined all successive utterances from the same speaker into one utterance. We then transformed each utterance into a binary vector with a value for each of the 14 LIWC categories (Pennebaker et al., 2007). We then formatted the utterances as a series of message-reply pairs, in which each utterance was treated as a reply to the previous utterance, and a message for the next utterance.

Analysis

For each pair of speakers A and B in a transcript, we computed four counts for each of the 14 LIWC categories: The number messages from A to B contained the category (N^{align}), the number of messages from A to B not containing the category (N^{base}), the number of replies containing the category to messages containing the category (C^{align}), and the number of replies containing the category to messages not containing the category (C^{base}). To generate robust parameter estimates, aggregated counts across all transcripts for the same parent and child into 6-mo. age bins, yielding 8 bins (youngest: 12–18 mo., oldest 54–60 mo.) When estimating α and β—the age-related scalars in the model, we numbered these bins from 1 to 8 and then subtracted the mean bin number from each (4.5). This centers the intercept (η^{align}) at the middle value, yielding the smallest average predictive error for other age bins.

We then fit the Hierarchical Alignment Model to the data separately for children and adults. Posterior distributions for all parameters were estimated using a Hamiltonian Monte Carlo sampler (Carpenter, 2015) with three independent chains, and 500 samples in each chain. The first 100 samples of each chain were discarded to ensure sufficient burnin based on inspection of trace plots that typically showed convergence after 50–75 samples. In addition, to provide a baseline for comparison, we also estimated alignment for the parent-parent interactions in the corpus.

Results and Discussion

The model's posterior parameter estimates are shown in Figure 2. Overall, alignment parameter estimates were above zero for both parents and children, suggesting that both groups aligned. Parents aligned reliably more to their children than children aligned to parents, and parents also aligned reliably more to their children than to each other. Thus, in the aggregate, we can conclude that parents coordinate to their children more than they coordinate to other adults.

The linguistic tuning hypothesis makes two predictions. First, parents might change the words they produce, using simpler words with younger children. If so, we would expect an increase in their likelihood of producing optional function words (positive β). In line with previous analyses, we find no evidence of this (Newport et al., 1977; Hayes & Ahrens, 1988). (We do, however, find a large and reliable increase in *children's* use of these words as they grow older, as would be expected based on their developing linguistic competence.)

Second, parents could use the same words, but be less contingent on children's production—repeating less, clarifying less, rephrasing less. This second prediction is precisely what we observe in both the aggregate parameter estimates and the developmental parameter estimates (Figure 3). Parental alignment decreases reliably over development, and children's alignment shows a similar trend. Thus, parent-child conversations become gradually less over-coordinated over development, settling to the adult-adult baseline as children need less scaffolding to be successful communicators.

In addition to estimating population-level parameters for alignment and fine-tuning, our hierarchical model also estimates parameters for each of the individual adults and children in CHILDES. Examining these parameters shows both

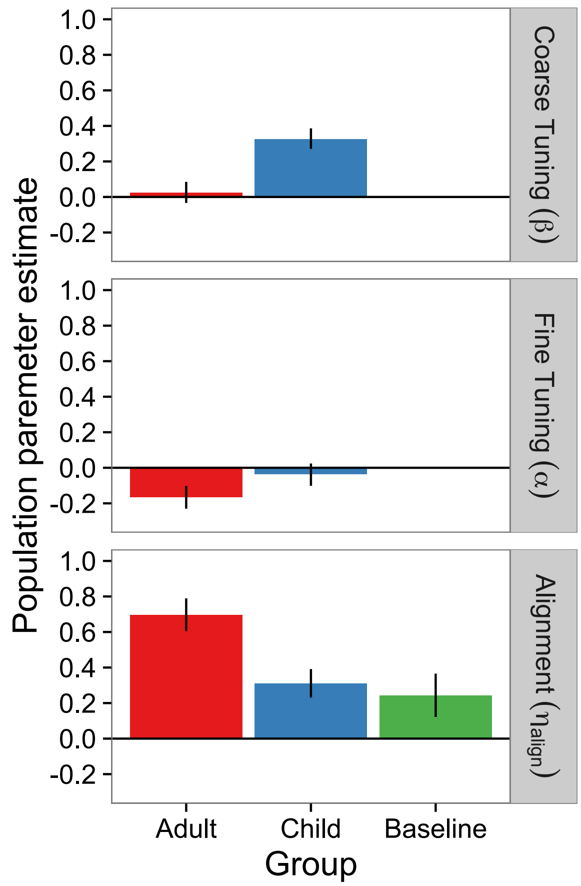

Figure 2: Posterior parameter estimates for population levels of alignment (η^{align}), fine tuning (α), and coarse tuning (β) both parents and children, as well estimated parent-parent alignment for a baseline. Bars indicate means, and error-bars indicate the 95% highest posterior density intervals.

consistency and variability across dyads (Figure 4). Across children, parental alignment is consistently highest early in development, but both children and parents vary in their level of alignment and in the rate at which it changes across development.

Repetition vs. Category stuff here? [mcf: YES!]

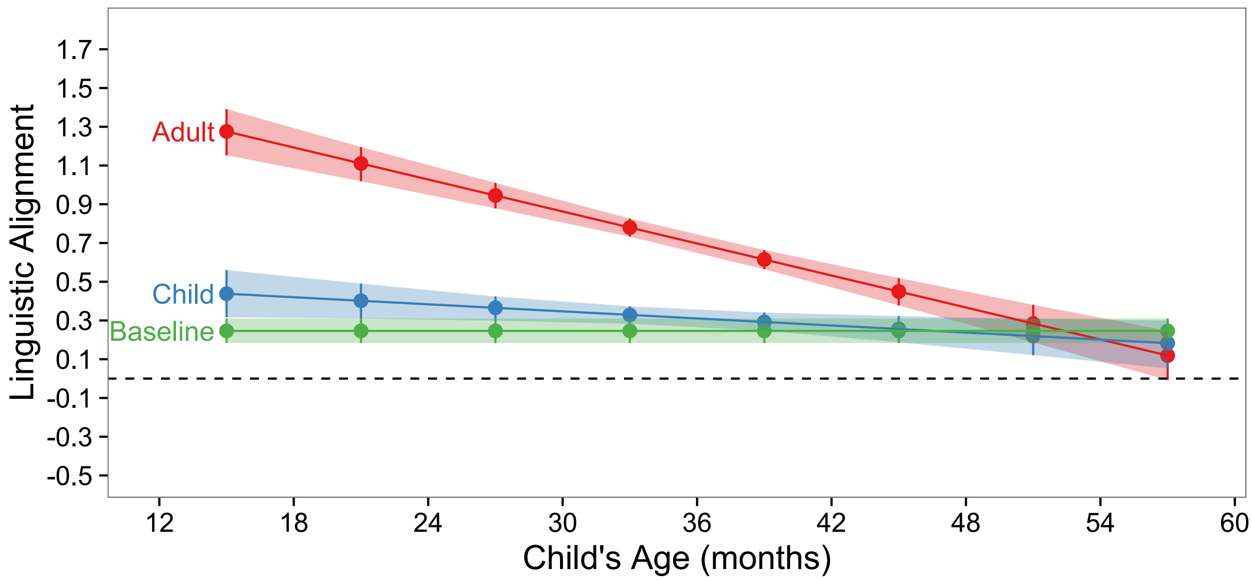

Figure 3: Model-estimated changes in linguistics alignment by 6-month-window. Over the course of development, both parents and children decrease in their linguistic alignment until they are indistinguishable from adult-adult interaction baselines. Points indicate the mean of the posterior distribution, lines and shaded regions indicate 68% highest probability density intervals, equivalent to one standard deviation.

General Discussion

Although even the youngest infants are equipped with the capacity to learn from the language they hear, their learning is highly constrained by their developing attentional and memory systems (Vlach & Johnson, 2013). How do children acquire language so rapidly despite these cognitive constraints? One hypothesis is that the the language they hear is tuned to their learning capacities, providing the right kind of information at the right time (Snow, 1972; Vygotsky, 1978).

Previous attempts to test this hypothesis have produced a mixture of evidence, with some in strong support of the linguistic tuning hypothesis (e.g. Hirsh-Pasek et al., 1984; Chouinard & Clark, 2003), others in strong opposition (e.g. Brown & Hanlon, 1970; Newport et al., 1977), and yet others finding mixed support and interesting individual differences (e.g. ?, ?, ?). Our work leverages the power of hierarchical Bayesian models to bring together data from two orders of magnitude more children than previous work.

We find at the population level that parents indeed provide linguistic input that is calibrated to children's development. The structural elements of their utterances are strongly contingent on children's previous productions early in development, but gradually approach adult-adult levels of coordination. In line with previous work, we do not find evidence of change at a global level: Parents do not use different words when talking to children of different ages (Hayes & Ahrens, 1988). Instead, calibration is local—sensitive to the on-going conversation and attuned to children's processing in context.

As opposed to across-the-board simplification, linguisic tuning may be particularly beneficial for learning. Because context provides so much support for children's comprehension, perhaps the best thing that parents can do is provide support in context. And critically, the linguisic tuning account does not require that parents consciously have as their goal the optimization of children's learning. They need not be teachers. If parents want to communicate with their children, and their children need significant linguistic support, they will have no choice to align.

Acknowledgments

We are grateful to Aaron Chuey and Jake Prasad for help with an earlier version of the model. This work was supported by NIH NRSA F32HD075577 to DY and NSF BCS 1528526 and 1456077 to MCF.

References

Brown, R., & Hanlon, C. (1970). Derivational complexity and order of acquisition in child speech. In *Cognition and the development of language* (Vol. 8). New York: Wiley.

Carpenter, B. (2015). Stan: A probabilistic programming language. *Journal of Statistical Software*.

Chouinard, M. M., & Clark, E. V. (2003). Adult reformulations of child errors as negative evidence. *Journal of child language*, *30*(03), 637–669.

Demuth, K., Culbertson, J., & Alter, J. (2006). Word-minimality, epenthesis and coda licensing in the early acquisition of english. *Language and Speech*, *49*(2), 137–173.

Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., & Tylén, K. (2012). Coming to terms quantifying the benefits of linguistic coordination. *Psychological science*, 0956797612436816.

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in cognitive sciences*, *8*(1), 8–11.

Giles, H., Coupland, J., & Coupland, N. (1991). *Contexts of accommodation: Developments in applied sociolinguistics*. Cambridge University Press.

Gogate, L. J., Bahrick, L. E., & Watson, J. D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child development*, *71*(4), 878–894.

Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*(2), 109–135.

Hayes, D. P., & Ahrens, M. G. (1988). Vocabulary simplification for children: A special case of motherese? *Journal of child language*, *15*(02), 395–410.

Hirsh-Pasek, K., Treiman, R., & Schneiderman, M. (1984). Brown & hanlon revisited: mothers' sensitivity to ungrammatical forms. *Journal of child language*, *11*(01), 81–88.

Ireland, M. E., Slatcher, R. B., Eastwick, P. W., Scissors, L. E., Finkel, E. J., & Pennebaker, J. W. (2011). Language style matching predicts relationship initiation and stability. *Psychological science*, *22*(1), 39–44.

Kacewicz, E., Pennebaker, J. W., Davis, M., Jeon, M., & Graesser, A. C. (2013). Pronoun use reflects standings in social hierarchies. *Journal of Language and Social Psychology*, 0261927X13502654.

Lieven, E., Salomo, D., & Tomasello, M. (2009). Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics*, *20*(3), 481–507.

Mayor, J., & Plunkett, K. (2011). A statistical estimate of infant and toddler vocabulary size from cdi analysis. *Developmental Science*, *14*(4), 769–785.

Montag, J. L., Jones, M. N., & Smith, L. B. (2015). The words children hear picture books and the statistics for language learning. *Psychological Science*, *26*(9), 1489–1496.

Montag, J. L., & MacDonald, M. C. (2015). Text exposure predicts spoken production of complex sentences in 8-and 12-year-old children and adults. *Journal of Experimental Psychology: General*, *144*(2), 447.

Newport, E., Gleitman, H., & Gleitman, L. (1977). Mother, i'd rather do it myself: Some effects and non-effects of maternal speech style. *Talking to children: Language input and acquisition*, 109–149.

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). Linguistic inquiry and word count: Liwc [computer software]. *Austin, TX: liwc. net*.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568.

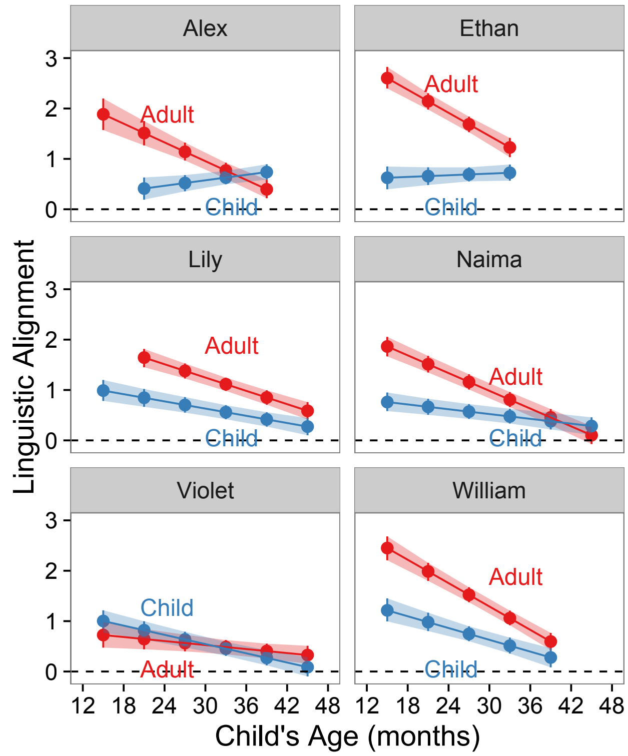

Figure 4: Model parameters.

Snow, C. E. (1972). Mothers' speech to children learning language. *Child development*, 549–565.

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, *7*(1), 53–71.

Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition*, *127*(3), 375–382.

Vygotsky, L. S. (1978). *Mind in society: The development of higher mental process.* Cambridge, MA: Harvard University Press.

Yuan, S., & Fisher, C. (2009). "really? she blicked the baby?" two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science*, *20*(5), 619–626.

Yurovsky, D., Yu, C., & Smith, L. B. (2012). Statistical speech segmentation and word learning in parallel: scaffolding from child-directed speech. *Frontiers in psychology*, *3*.