Research Article

The Emergence of an Abstract Grammatical Category in Children's Early Speech

Stephan C. Meylan[1], Michael C. Frank[2], Brandon C. Roy[2,3], and Roger Levy[4]

[1]Department of Psychology, University of California, Berkeley; [2]Department of Psychology, Stanford University; [3]The Media Lab, Massachusetts Institute of Technology; and [4]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology

Abstract

How do children begin to use language to say things they have never heard before? The origins of linguistic productivity have been a subject of heated debate: Whereas generativist accounts posit that children's early language reflects the presence of syntactic abstractions, constructivist approaches instead emphasize gradual generalization derived from frequently heard forms. In the present research, we developed a Bayesian statistical model that measures the degree of abstraction implicit in children's early use of the determiners "a" and "the." Our work revealed that many previously used corpora are too small to allow researchers to judge between these theoretical positions. However, several data sets, including the Speechome corpus—a new ultra-dense data set for one child—showed evidence of low initial levels of productivity and higher levels later in development. These findings are consistent with the hypothesis that children lack rich grammatical knowledge at the outset of language learning but rapidly begin to generalize on the basis of structural regularities in their input.

One of the most astonishing parts of children's language acquisition is the emergence of the ability to say and understand things that they have never heard before. This ability, known as *productivity*, is a hallmark of human language (Hockett, 1959; von Humboldt, 1836/1999). Adults' linguistic representations are almost universally described in terms of syntactic abstractions such as "determiner," "verb," and "noun phrase" (e.g., Chomsky, 1981; Sag, Wasow, & Bender, 1999). But do these same adultlike abstractions guide how children produce and comprehend language?

Some researchers have suggested a *generativist* view of syntactic acquisition: Adultlike abstractions guide children's comprehension and production from as early as they can be measured (Pinker, 1984; Valian, 1986; Yang, 2013). Other researchers have argued that adultlike syntactic categories— or at least their guiding role in behavior—emerge gradually with the accumulation of experience. According to such *constructivist* views, children's representations progress over time from memorized multiword expressions to specific item-based constructions and eventually generalize to abstract combinatorial rules (Braine, 1976; Pine & Lieven, 1997; Pine & Martindale, 1996; Tomasello, 2003).

Here, we focus on a key case study for this debate: the emergence of the capacity in English to produce a noun phrase (NP) by combining a determiner (Det, such as "the" or "a") with a noun (N). This capacity is exemplified in adult English by the context-free rule NP → Det N. Because of this knowledge, when adult native English speakers hear a novel count noun with "a" (e.g., "a blicket"), they know that combining the novel noun with "the" will also produce a permissible noun phrase (e.g., "the blicket"). Adultlike usage of this part of English syntax requires knowledge of the category "Det," the category "N," and the rule specifying how they are combined.

Corresponding Author:
Stephan C. Meylan, University of California, Berkeley, Department of Psychology, 3210 Tolman Hall, Berkeley, CA 94720
E-mail: smeylan@berkeley.edu

In recent years, this case study—noun-phrase productivity, with a focus on the use of determiners—has played an increasingly prominent role in the generativist-constructivist debate on children's language acquisition (Pine & Lieven, 1997; Pine & Martindale, 1996; Valian, 1986; Valian, Solt, & Stewart, 2009; Pine, Freudenthal, Krajewski, & Gobet, 2013; Yang, 2013). Whereas nouns often have referents in the child's environment, the semantic contribution of determiners to utterance meaning is more subtle (Fenson et al., 1994; Tardif et al., 2008). Thus one might expect determiners to be learned late (Valian et al., 2009). Yet children produce them relatively early, and their uses are overwhelmingly correct by the standards of adult grammar. Is this because children deploy adultlike syntactic knowledge? Or is it because they memorize and reuse specific noun phrases, which creates the illusion of full productivity?

Experimental methods have been of limited utility in resolving this question. Tomasello and Olguin (1993) found evidence for the presence of a nounlike productive object-word category in children between 20 and 26 months, presenting objects with nonce labels and eliciting reuse in novel syntactic contexts and morphological forms (using a "wug" test; Berko, 1958). But these data do not resolve the extent to which syntactic abstractions guide children's everyday speech. Instead, most work on early syntactic productivity has relied on observational language samples (Pine et al., 2013; Pine & Lieven, 1997; Pine & Martindale, 1996; Valian, 1986; Valian et al., 2009; Yang, 2013).

Making inferences about children's knowledge from observational evidence is difficult for a number of reasons, however. First, individual corpora of children's language have typically been small—consisting of weekly or monthly recordings of only a couple of hours. Second, nouns (like other words) follow a Zipfian frequency distribution (Zipf, 1935), in which a small number of words are heard often but most are heard only a handful of times. As a result, evidence regarding the range of syntactic contexts in which a given child uses an individual noun is weak for most nouns (Yang, 2013). These inferential challenges are sufficiently severe that, within the past several years, researchers on opposing sides of the productivity debate have drawn opposite conclusions from similar data sets (Pine et al., 2013; Yang, 2013). Making progress on children's syntactic productivity requires overcoming these challenges.

Here, we present a new, model-based framework for drawing inferences about syntactic productivity, which differs from previous methods in two critical respects. First, previous approaches assessed productivity via a summary statistic, the *overlap score*, computed from a sample of child language. This statistic is difficult to interpret because it may be biased by the size and composition of

the sample (as discussed in the next section). Here, in contrast, we treated productivity as one feature of a model of child language whose parameters can be estimated from a sample and whose overall fit to the data can be assessed. Second, we explicitly modeled item-based memorization and reuse of specific determiner-noun pairs from caregiver speech in the child's environment as an additional contributor to children's language production alongside syntactic productivity. Our framework encodes a continuum of hypotheses ranging between fully productive and fully item-based, and it allows us to assess how children at any given point in development balance these two knowledge sources in their production of determiner-noun combinations.

We applied this model to a wide range of longitudinal corpora of children's speech, including the Speechome corpus (Roy, Frank, DeCamp, Miller, & Roy, 2015), a new high-density set of recordings of one child's early input and productions. Our model revealed that many of the conventional corpora analyzed in previous research (Pine et al., 2013; Pine & Lieven, 1997; Pine & Martindale, 1996; Valian, 1986; Valian et al., 2009) are too small to enable researchers to draw high-confidence inferences. An exploratory analysis of the Speechome data, the coverage of which is both denser and from earlier in development, provided evidence for low initial levels of productivity followed by a rapid increase starting around 24 months. Several other data sets provided corroboratory evidence. Contra full-productivity accounts, our model showed that syntactic productivity is very low in the first months of determiner use in these data sets. At the same time, the current work constrained the timeline of constructivist accounts. We found a rapid early increase in productivity—in the child from the Speechome corpus, this increase occurred within a few months of the onset of combinatorial speech, prior to the beginning of many of the data sets that have been used previously to address this question. We conclude by discussing the need for denser data sets to provide conclusive evidence on questions about the roots of syntactic abstraction.

Previous Work and Present Goals

Previous investigations have focused on the overlap score, a summary statistic of productivity (Pine et al., 2013; Pine & Lieven, 1997; Pine & Martindale, 1996). Overlap is calculated from the distribution of determiner-noun pairings in a sample, as the proportion of nouns that appear with both "a" and "the" out of the total number of nouns used with either. While initial investigations suggested that young children use comparatively fewer nouns with both determiners than with just one (Pine & Lieven, 1997; Pine & Martindale, 1996), overlap scores are highly dependent on sample size because of the Zipfian

distribution of noun frequencies (Valian et al., 2009; Yang, 2013). In addition, this statistic is not well-suited for distinguishing true changes in grammatical productivity from increased exposure to the relevant words (e.g., hearing both "the dog" and "a dog" independently and subsequently repeating these, even without abstraction; Valian et al., 2009; Yang, 2013).

Two recent investigations used more sophisticated techniques to address issues of sample size. Yang (2013) constructed a null-hypothesis full-productivity model in which each noun had the same distribution over determiner pairings (no item-specific preferences) and showed that it predicted overlap scores well for six children in the Child Language Data Exchange System (CHILDES) database (MacWhinney, 2000). Pine et al. (2013), in contrast, developed a noun-controlled method for comparing adults' and children's productivity scores in a given sample and rejected a full-productivity null hypothesis. Neither of these methods, however, is well suited to tracking developmental changes in productivity because of their focus on the overlap score. If item-based knowledge plays a role in children's productions, overlap might increase over time even without any changes in productivity, simply because children have heard more determiner-noun pairs.

Here, we took a fundamentally different approach from previous work to address the challenge of decoupling genuinely productive behavior from what might be expected on the basis of experience. We proceeded from the observation that there are two sources of information by which a speaker could know that a particular determiner-noun pair belongs to English and thus potentially produce it: (a) direct experience with that specific determiner-noun pair and (b) a productive inference using knowledge abstracted from experience with different determiner-noun pairs (and perhaps other input). Measuring a given speaker's productivity from corpus data requires assessing the extent to which the speaker's language use reflects productivity above and beyond what can be attributed to direct experience.

We defined a probabilistic model of determiner-noun production that considers both knowledge sources. In our model, the contribution of productive knowledge can range along a continuum from none (i.e., the child is capable only of imitating caregiver input—an idealized version of an *island learner* as described in Tomasello, 1992) to extreme (i.e., the child is a *total generalizer*; note that such a model is equivalent to the null-hypothesis model in Yang, 2013). Specific model parameters corresponded to the contributions of these two information sources, and we used Bayesian inference to estimate likely values of these parameters for a corpus sample given both the child's determiner-noun productions and caregiver input. By comparing temporally successive

samples for a given child, we used this model to infer the child's change in syntactic productivity over time. Because our model is fully Bayesian, we were also able to estimate the level of certainty in our estimates, which critically allowed us to avoid overly confident inferences when data were too sparse.

Method

Model

We modeled the use of each noun token with a specific determiner as the output of a probabilistic generative process. We assumed that each noun has its own determiner preference ranging from 0 (a noun used only with "a") to 1 (a noun used only with "the"). We then explicitly modeled cross-noun variability by assuming some underlying distribution of determiner preferences across all nouns. Lower cross-noun variability indicates that nouns behave in a more classlike fashion, while higher variability indicates little generalization of determiner use across nouns.

Formally, each noun type can be thought of as a coin whose weight corresponds to its determiner preference. Each use of that noun type with a determiner is thus analogous to the flip of that weighted coin, where heads indicates the use of the definite determiner, and tails corresponds to the indefinite. A sequence of noun uses are thus draws from a binomial distribution with success parameter μ corresponding to the determiner preference. The determiner preference for each noun is drawn from a beta distribution with mean μ_0 (the underlying "average" preference across all nouns) and scale v, yielding a *hierarchical beta-binomial* model (Gelman, Carlin, Stern, & Rubin, 2003).[1]

Under this model, a child's determiner productions for each noun he or she uses are guided by a combination of the two information sources mentioned previously— (a) direct experience and (b) productive knowledge— and the strength of each information source's contribution to the child's productions is determined by a weighting parameter. For direct experience, a parameter η determines how effectively the child learns from noun-specific determiner productions in his or her linguistic input; for productive knowledge, a parameter v determines how strongly the child applies productive knowledge of determiner use across all nouns. These parameters η and v do not trade off against each other but rather play complementary roles in accounting for a child's productions: As η increases, the variability across nouns in a child's determiner productions can more closely match the variability in his or her input, while as v increases, the child is increasingly able to produce determiner-noun pairs for which he or she has not received sufficient evidence

from caregiver input. (For more details, see Parameters of the Beta-Binomial Model in the Supplemental Material available online; our complete hierarchical Bayesian model and variable definitions are presented in Fig. S1 in the Supplemental Material.)

Because we lacked exhaustive recordings of caregiver input, we treated unrecorded caregiver input as a latent variable drawn from the same distribution as aggregated caregiver input and inferred it jointly with model parameters (see Details of the Imputation in the Supplemental Material). The theoretically critical target of inference was ν, the strength of the child's productive knowledge of determiner-noun combinatorial potential, which could range from 0 (an extreme island learner whose determiner preference for a given noun is guided exclusively by his or her direct experience with that noun and whose noun-specific determiner preferences are likely to be skewed toward 0 or 1) to near infinity (an extreme overgeneralizer who has identical determiner preference for all nouns, Fig. 1a).

We used Markov-chain Monte Carlo sampling to infer confidence intervals over η, μ_0, and ν from a child's recorded productions and linguistic input. But a single recording of a child typically does not yield high confidence in these estimates because of the relatively low numbers of productions for individual nouns. To overcome this issue, we used two different methods for constructing sufficiently large samples of child and caregiver tokens to evaluate the developmental trajectory of the ν parameter: split-half and sliding-window analyses.

First, in the *split-half* analysis, we divided the data for each child into distinct early and late time windows with an equal number of tokens, denoted with the subscripts t_1 and t_2. Separate parameter sets (μ, ν, η) were inferred for the first and second windows; for a given sample from the joint posterior, the changes in parameters from the first window to the second were calculated as follows:

$$\Delta \nu = \nu_{t_2} - \nu_{t_1}$$

$$\Delta \mu = \mu_{t_2} - \mu_{t_1}$$

$$\Delta \eta = \eta_{t_2} - \eta_{t_1}$$

These variables may be treated as targets of inference, over which highest-posterior-density (HPD) intervals may be computed. Our principal target of inference was $\Delta \nu$, the change in the contribution of productive knowledge to the child's determiner use. This two-window approach maximized statistical power, but did so at the expense of a detailed time-related trajectory: For children with longer periods of coverage, this estimate may have grouped together several distinct developmental time periods.

Second, as an exploratory technique, we also used our model to measure finer-grained changes in parameter estimates across development via a *sliding-window* approach, in which the model was fitted to successively later subsections of the corpus of child productions. Each window also included the corresponding adult productions that occurred prior to or during that subsection. In this case, we fitted the model to successive 1,024 token windows of the child's speech, advancing by 256 tokens for each sample. This method yielded a higher-resolution time course than the split-half analysis, though at the expense of less-constrained parameter estimates, especially for the smaller corpora. (For more details on inferring model parameters, see Model Fitting Procedure in the Supplemental Material.)

Our approach is an example of Bayesian data analysis (Gelman et al., 2003). We created a cognitively interpretable model that captured the spectrum of different hypotheses, from item-based learning to full productivity. We could then infer, for a particular data set, where on the spectrum the data fall. In a classic predictive model, parameters are fitted—or overfitted—to some external performance standard. In contrast, our model summarized a particular aspect of the data set and gave an estimate of the relative certainty of this summary measurement.

Data

We used a large set of publicly available longitudinal developmental corpora of recordings of children and their caregivers from the CHILDES archive (MacWhinney, 2000). Four of these corpora have been examined previously for early evidence of grammatical productivity: the Providence corpus (Demuth, Culbertson, & Alter, 2006), the Manchester corpus (Theakston, Lieven, Pine, & Rowland, 2001), the Brown corpus (Brown, 1973), and the Sachs corpus (Sachs, 1983). We additionally analyzed four single-child corpora: Bloom (Bloom, Hood, & Lightbown, 1974), Kuczaj (Kuczaj, 1977), Suppes (Suppes, 1974), and Thomas (Lieven, Salomo, & Tomasello, 2009). These eight corpora yielded usable data for a total of 26 children (see Table S1 in the Supplemental Material for the names, age ranges, and other information about all children from the analyzed corpora). While high-density data with rich annotations exist for all of these corpora, coverage started in most cases well after the onset of combinatorial speech and is sparse under 2 years of age, the time interval necessary for characterizing initial levels of grammatical productivity.

To address these shortcomings, we additionally analyzed the densest longitudinal developmental corpus in existence, the Speechome corpus (Roy et al., 2015). The Speechome corpus covers the period 9 through 25

a
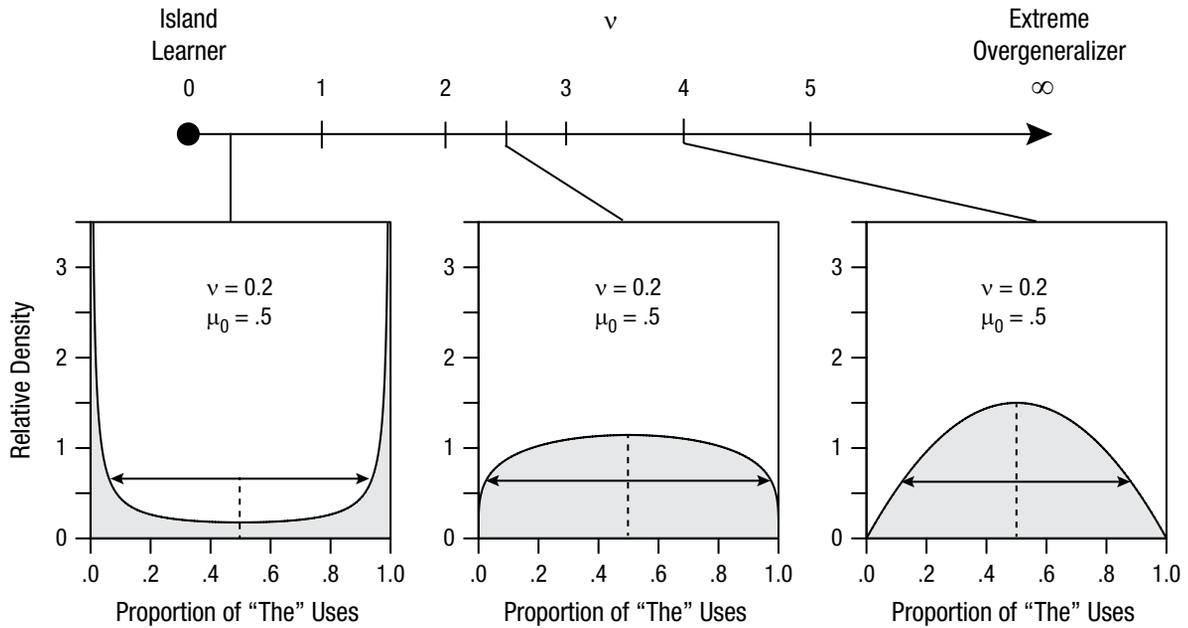
b
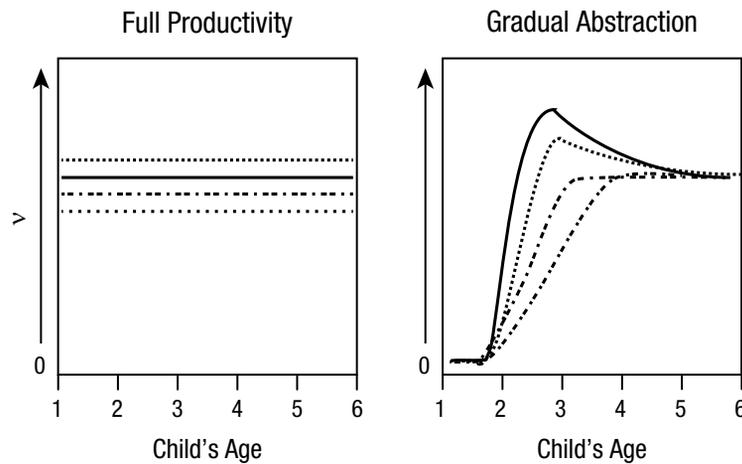
Fig. 1. Interpretation of the ν parameter and expected trajectories of grammatical productivity under two accounts of early language acquisition. An illustration of ν, the strength of the child's productive knowledge of determiner-noun combinatorial potential, is shown in (a). At low values of ν, little or no information is shared between nouns. At higher values of ν, nouns exhibit more consistent usage as a class, which indicates the existence of a productive rule governing the combination of determiners and nouns. The graphs show the distribution of the mean proportion of uses of "the," separately for low, medium, and high levels of ν; μ_0 represents the mean proportion of definite determiner usage across nouns, set here at .5 in all three graphs. Schematized trajectories for the development of grammatical productivity (b) are shown according to two competing theories. The full-productivity account holds that children use determiners productively as early they can produce them and that this productivity changes little across development. The gradual-abstraction account, in contrast, posits that children are initially nonproductive and that productivity increases over the first several years of development. For each account, four hypothetical trajectories are shown.

months in the life of a single child, and it contains video and audio recordings of nearly 70% of the child's waking hours, with transcripts for a substantial portion of these (Vosoughi & Roy, 2012). While transcription of the Speechome corpus is a work in progress, the version used here contains approximately 4,300 noun phrases

with articles produced by the child before 25 months of age and includes dense coverage of child-accessible caregiver speech, with some 196,300 noun phrases in the same time period. The Speechome corpus supports more detailed inferences about the developmental time course in the second year of life. The Speechome corpus is also
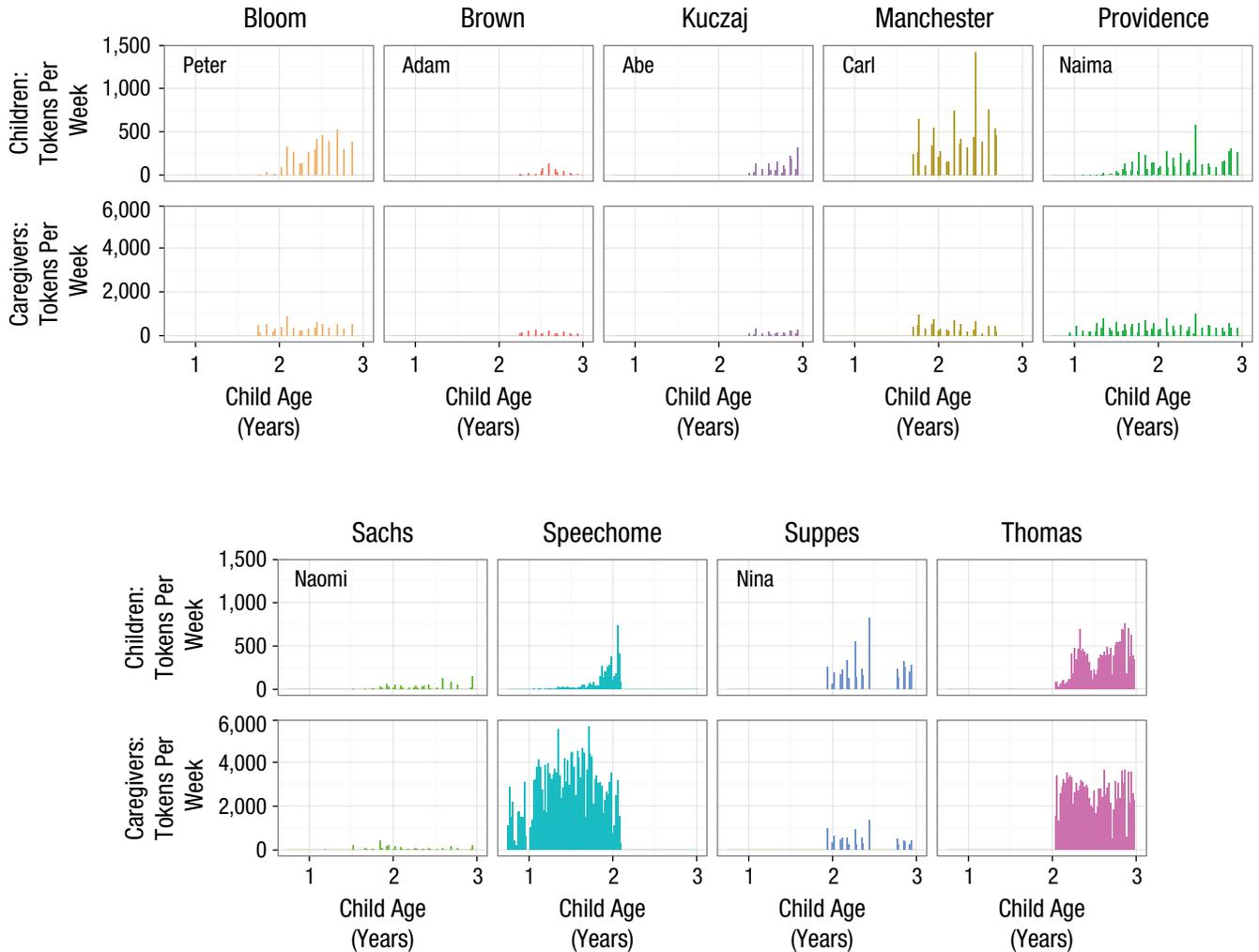
Fig. 2. Number of recorded determiner-noun pairs used per week by children prior to age 3 years and by their corresponding caregivers, separately for each of the nine corpora analyzed. For multichild corpora (Providence, Brown, Manchester, and Sachs), values are shown only for the child with the most determiner-noun pairs; the other corpora contained data from only one child. Note that the number of weekly observations from children and caregivers are presented on differing scales (0–1,500 and 0–6,000, respectively).

distinguished in the quantity of child-available adult speech, with nearly 80% more caregiver tokens than in the corpus for the next-best-represented child, Thomas. Figure 2 shows comparative densities for adult and child determiner-noun pairs for the child with the most data in each corpus.[2]

We assessed our model using seven different methods for extracting determiner-noun data from each corpus. These data treatments reflect a range of assumptions regarding the availability of phrase structure for identifying which noun corresponds to each determiner, whether information can be shared between morphologically inflected forms, and whether the child is considering only singular forms in the language. In the absence of reliable morphological tags, the Thomas and Speechome corpora were assessed on four data treatments each. (For additional technical details, refer

to Data Preparation in the Supplemental Material. We have placed model code, noun-anonymized Speechome data, and auxiliary code necessary to reproduce our research in a public GitHub repository at https://github.com/smeylan/determiner_learning.)

Results

The two hypotheses represented in the literature—full productivity or gradual abstraction over item-based knowledge (Fig. 1b)—make contrasting predictions regarding initial productivity and the effects of developmental change. The full-productivity account predicts a nonzero initial level combined with a negligible effect of developmental time—productivity does not increase with exposure to more data. The account positing gradual abstraction over item-based knowledge, in contrast, predicts near-zero
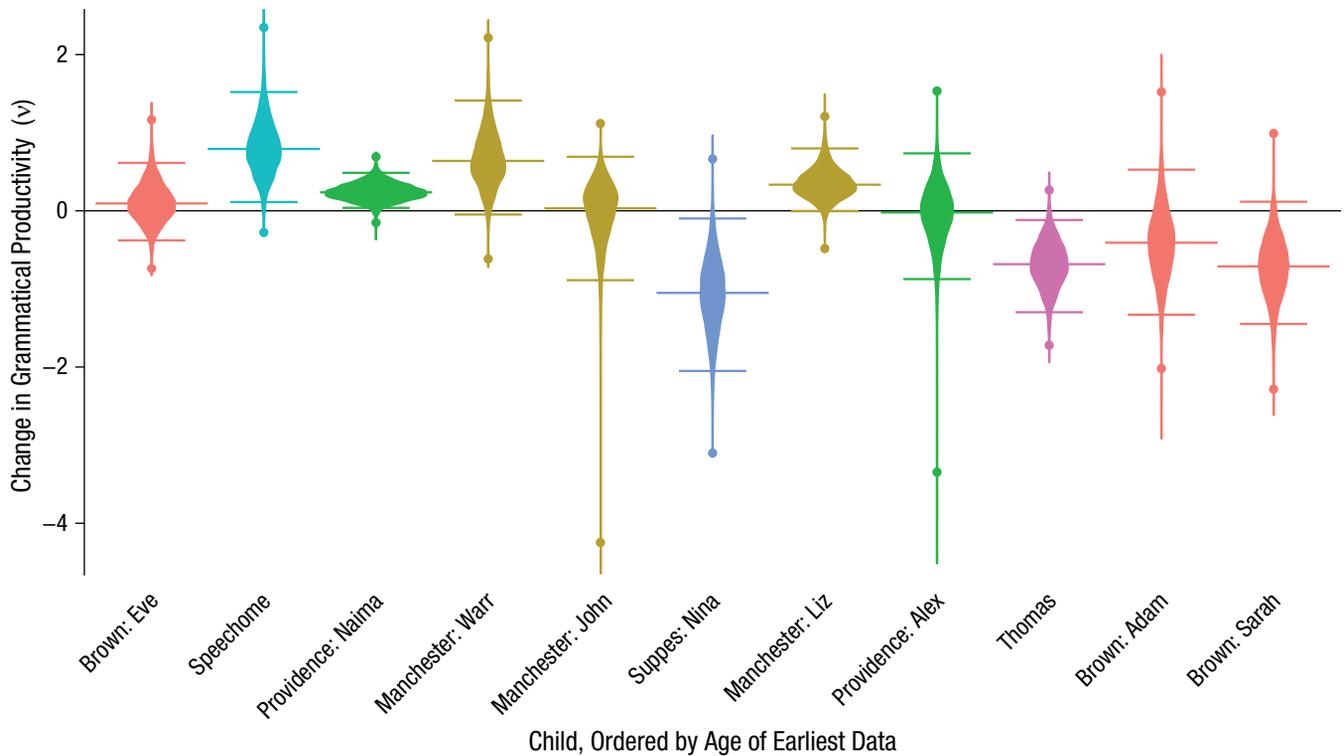
Fig. 3. Posterior estimates for change in grammatical productivity between the first and second half of children's corpora (Δv) for 11 of the children from eight of the nine corpora. The longest horizontal lines indicate the median of the posterior, and the shorter horizontal lines the 95% interval of highest posterior density (HPD). Points indicate the 99.9% HPD interval. Estimates for the remainder of the children ($n = 16$) are not displayed because their data had poorly constrained posteriors (i.e., the 99.9% HPD for v fell outside the interval [0,3] for at least one of the time periods).

initial productivity, which would indicate the absence of syntactic-category knowledge in the earliest productions, and a positive relationship with developmental time corresponding to the gradual induction of abstract categories throughout childhood.

Split-half method

To test for changes in productivity, we assessed the null hypothesis that 0 (no change) was within the 99.9% HPD interval for the posterior estimate of Δv, the difference in v estimates between the first and second half of tokens for each child. (We used the 99.9% criterion because of the large number of independent comparisons implied by this analysis—one for each of the 27 children.) By this standard, only one child (from the Speechome corpus, in three of four data treatments) showed a significant increase in productivity (Figs. 3 and 4). The remaining data treatment for the Speechome corpus was strictly positive within the 95% HPD interval. For three other children—Liz from the Manchester corpus (in three of seven data treatments), Naima from the Providence corpus (one of seven treatments), and Warr from the Manchester corpus (one of seven treatments)—change was strictly positive within the 95% HPD interval for at least one of the data treatments.

These findings suggest some early increases in productivity. (Results for all seven data treatments are presented in Fig. S4 in the Supplemental Material.)

We also found apparent *decreases* in grammatical productivity for several of the older children. Thomas (two of four treatments within the 99.9% HPD interval, one in the 95% HPD interval), Sarah from the Brown corpus (one in the 99.9% and four in the 95% HPD interval), and Nina from the Suppes corpus (two in the 99.9% and two in the 95% HPD interval) showed strictly negative changes. The timing of these decreases was consistent with a phase of overregularization, during which they were more willing to use determiner-noun combinations that are rare or unattested in adult speech, such as *a sky*, followed by a decrease toward adultlike levels. Consistent with this hypothesis, results showed that increases in v tended to occur in data sets from younger children ($p = .009$ by rank-sum test on the data shown in Fig. 3). Together, these results are broadly consistent with constructivist hypotheses, in that we found minimal evidence of productivity in the earliest multiword utterance coupled with a development-related increase in productivity soon thereafter. However, our results deviated slightly from the proposal of gradual emergence of abstract schema from item-specific exemplars, as set forth in
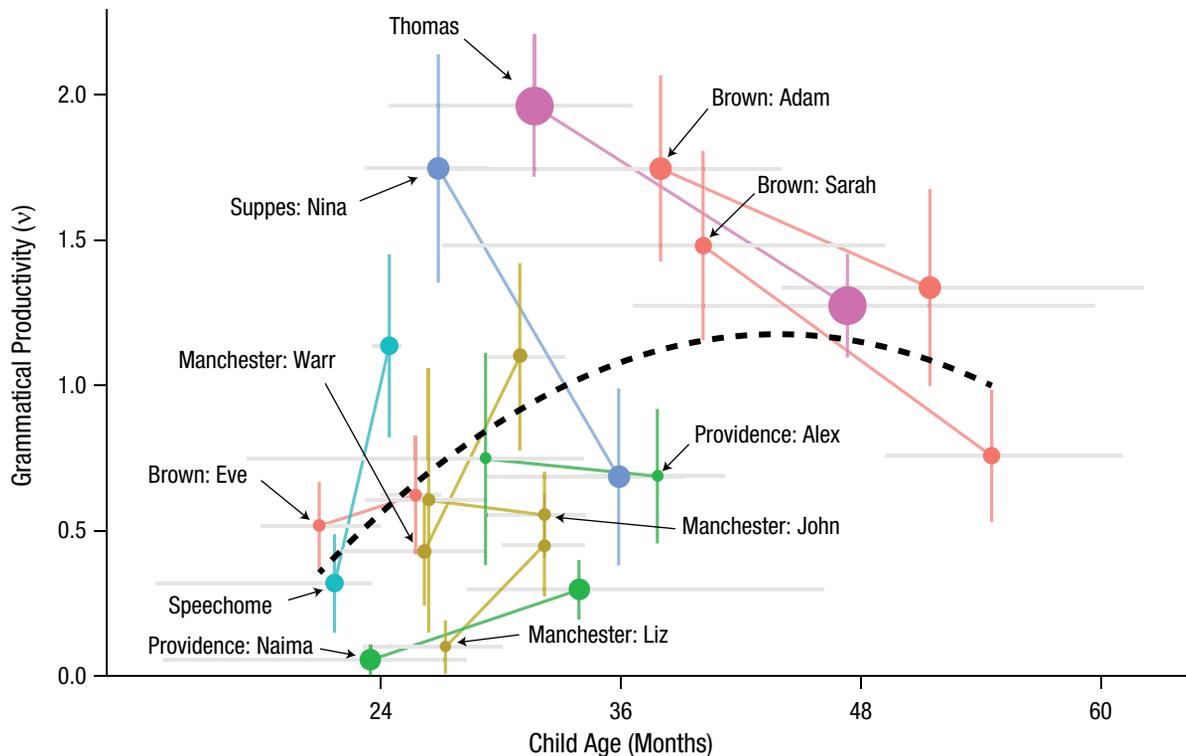
Fig. 4. Inferred developmental trajectory for determiner productivity, ν, for 11 of the children from the nine corpora, plotted by age in months. Marker size corresponds to the number of child tokens used for each child. Horizontal lines indicate the temporal extent of the tokens used to parameterize the model at each point; vertical lines indicate the standard deviation of the posterior. The best-fitting quadratic trend is indicated by the dashed black line. Different corpora are represented by different colors. Estimates for the remainder of the children ($n = 16$) are not displayed because their data had poorly constrained posteriors (i.e., the 99.9% highest posterior density for ν fell outside the interval [0,3] for at least one of the time periods).

Abbot-Smith & Tomasello (2006). The possibility of a decrease in determiner productivity later in development suggests that while children may construct abstract generalizations from their input, they may also use input later in development to constrain overly general abstract schema (along the lines schematized in the top two trajectories shown in Fig. 1b, right panel).

Our model was defined independently from overlap scores, the primary measure of productivity used in previous literature. We took advantage of this independence to use overlap as a model-validation method. Although a simple overlap measure is not useful for characterizing productivity and comparing across children, we used it to validate our model within individuals. We did this by sampling new simulated determiner productions from the fitted model's distribution on child determiners for each time window, computing overlap, and then comparing the results to the empirical values from that same child. Empirical overlap fell within the 95% range of simulated overlap scores for all children, which validated the model's overall fit to the data. (For additional details, see Results in the Supplemental Material.)

Sliding-window method

The higher temporal resolution of the sliding-window method revealed changes in grammatical productivity consistent with the split-half analysis, with major increases in productivity for the child in the Speechome corpus and Warr and major decreases for Thomas, Adam, and Sarah (Fig. 5, left column). The sliding-window models also revealed significant variability in ν not related to age (e.g., Naima from the Providence corpus). Children showed highly variable preference for the definite versus indefinite determiner (μ) over early development (Fig. 5, center column). In addition, using the same validation technique described previously, we found that simulated overlap from sliding-window estimates was strongly correlated with empirical values (Pearson's *r*s of .940–.951 across data treatments; Fig. 5, right column).

Discussion

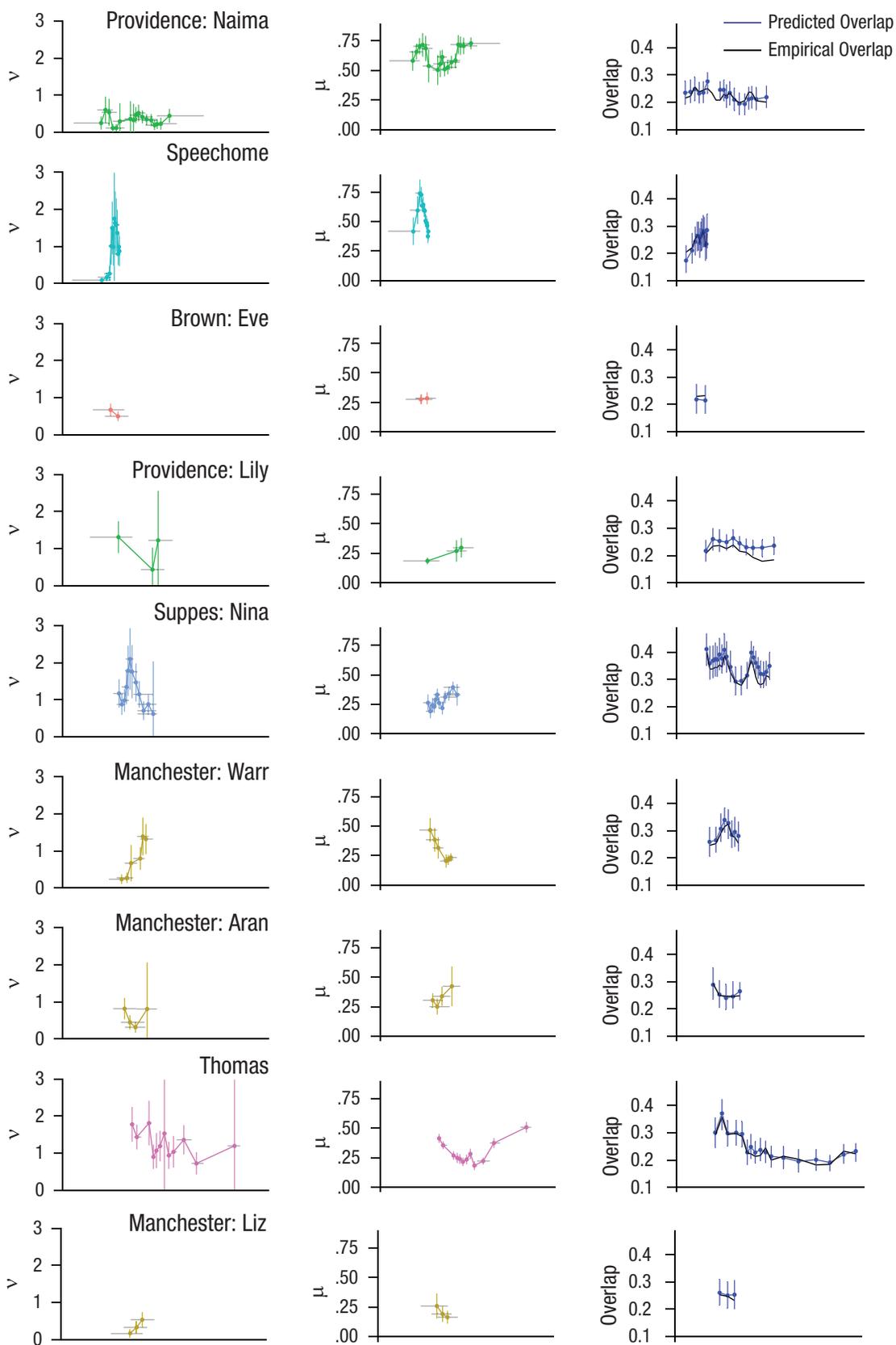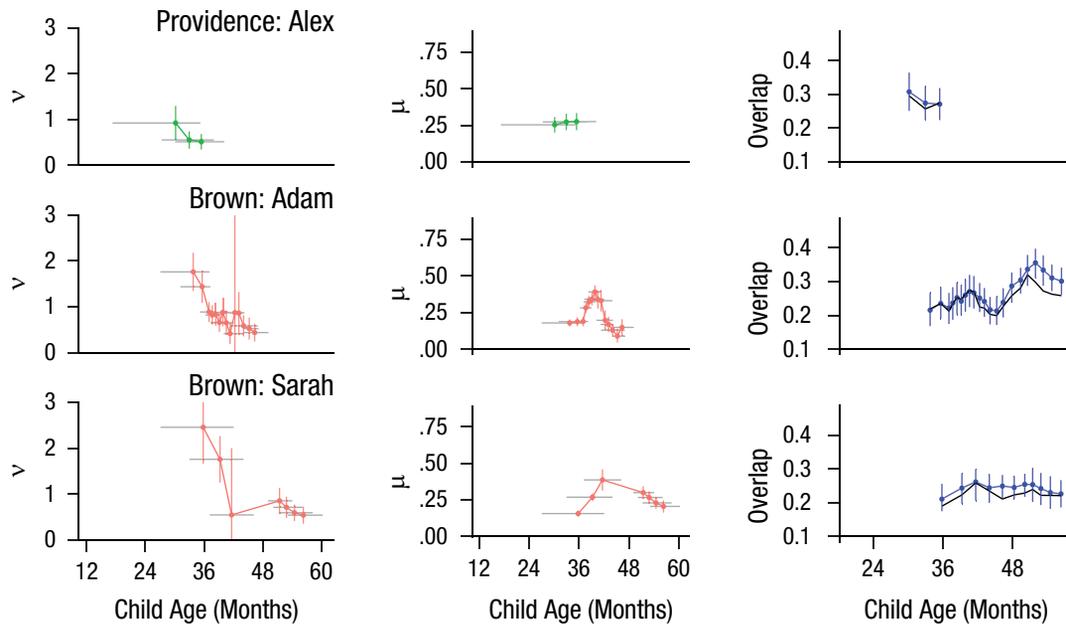The model-based statistical approach presented here for analyzing child language is the first method that allows

Fig. 5. *(continued on next page)*

Fig. 5. Determiner productivity (ν), mean determiner preference (μ), and predicted and empirical overlap scores for the 11 children included in the split-half analysis. Vertical lines show the 99% interval of highest posterior density for ν, μ, and overlap predicted by the current model. Horizontal lines indicate the temporal extent of the tokens used to fit the model at each point.

the respective contributions of productivity and item-based knowledge to be teased apart. Our analysis revealed two key findings. First, children's syntactic productivity changes over development. Several of the youngest children showed increases in productivity, with the strongest evidence in the largest data set, Speechome. In addition, some older children showed decreases in productivity. This trend might suggest a period of particularly strong generalization followed by a retreat, similar to the pattern observed in morphological domains (e.g., Pinker, 1991; Rumelhart & McClelland, 1985), as well as verb-argument structure (Ambridge, Pine, & Rowland, 2011; Bowerman, 1988).

Second, for the majority of children, our model placed wide confidence intervals on productivity estimates, which indicates that the available data were likely not sufficient to yield precise developmental conclusions. The data for these children typically included a maximum of 1 hr per week of transcripts; furthermore, most of the children's productions in these data sets were collected after each child's second birthday. If adultlike categories are constructed early—soon after the onset of word combination—many of these data sets begin too late to provide decisive evidence regarding the trajectory of early development. The trend line shown in Figure 4 is suggestive rather than conclusive; additional data sets would be required to test whether the pattern is robust within the developmental trajectory of a single child. These results underscore the critical importance of dense,

naturalistic data for understanding the development of linguistic knowledge in early childhood.

Debates about the emergence of syntactic productivity have typically oscillated between two poles: immediate, full productivity early in development or accumulation of item-specific knowledge with gradually increasing levels of productivity. Our approach parameterizes the space of models between these poles. In the future, it can be adapted to characterize productivity in other simple morphosyntactic phenomena and in other languages. In the key case study of English determiner productivity, applying our model to new, dense data yielded support for constructivist accounts and further constrained the developmental timeline within these accounts. While children's earliest multiword utterances may be islandlike, grammatical productivity emerges rapidly thereafter.

Action Editor

Matthew A. Goldrick served as action editor for this article.

Author Contributions

All authors designed the research. S. C. Meylan and B. C. Roy conducted the research. S. C. Meylan, M. C. Frank, and R. Levy analyzed the results. All authors wrote the manuscript.

Acknowledgments

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797616677753

Open Practices

All data are publicly available via TalkBank and can be accessed at http://childes.talkbank.org/. The complete Open Practices Disclosure for this article can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797616677753. This article has received the badge for Open Data. More information about the Open Practices badges can be found at https://osf.io/tvyxz/wiki/1.%20View%20the%20Badges/ and http://pss.sagepub.com/content/25/1/3.full.

Notes

1. Many readers may be more familiar with the more common parameterization of the beta distribution in terms of the following shape parameters: $\alpha = \mu\nu$ and $\beta = (1 - \mu)\nu$.

2. Data sets that capture large amounts of an individual family's experience, such as Speechome, pose unique privacy risks. In order to share reproducible data while maintaining privacy, we present determiner-noun count data from the Speechome corpus while obfuscating the identities of the specific nouns the child produced.

References

Abbot-Smith, K., & Tomasello, M. (2006). Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review*, *23*, 275–290.

Ambridge, B., Pine, J., & Rowland, C. (2011). Children use verb semantics to retreat from overgeneralization errors: A novel verb grammaticality judgment study. *Cognitive Linguistics*, *22*, 303–323.

Berko, J. (1958). The child's learning of English morphology. *Word*, *14*, 150–177.

Bloom, L., Hood, L., & Lightbown, P. (1974). Imitation in language development: If, when, and why. *Cognitive Psychology*, *6*, 380–420.

Bowerman, M. (1988). The 'no negative evidence' problem: How do children avoid constructing an overly general grammar? In J. Hawkins (Ed.), *Explaining language universals* (pp. 73–101). Oxford, England: Basil Blackwell.

Braine, M. D. S. (1976). Children's first word combinations. *Monographs of the Society for Research in Child Development*, *41*(1, Serial No. 164), 1–97.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, The Netherlands: Foris Publishers.

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

Fenson, L., Dale, P., Reznick, J., Bates, E., Thal, D., Pethick, S., . . . Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5, Serial No. 242), 1–185.

Gelman, A., Carlin, J., Stern, H., & Rubin, D. (2003). *Bayesian data analysis* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC.

Hockett, C. (1959). Animal languages and human languages. *Human Biology*, *31*, 32–39.

Kuczaj, S. (1977). The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior*, *16*, 589–600.

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*. Mahwah, NJ: Erlbaum.

Pine, J., Freudenthal, D., Krajewski, G., & Gobet, F. (2013). Do young children have adult-like syntactic categories? Zipf's law and the case of the determiner. *Cognition*, *127*, 345–360.

Pine, J., & Lieven, E. (1997). Slot and frame patterns and the development of the determiner category. *Applied Psycholinguistics*, *18*, 123–138.

Pine, J., & Martindale, H. (1996). Syntactic categories in the speech of young children: The case of the determiner. *Journal of Child Language*, *23*, 369–395. doi:10.1037/0012-1649.22.4.562

Pinker, S. (1984). *Language learnability and language development*. Cambridge, England: Cambridge University Press.

Pinker, S. (1991). Rules of language. *Science*, *253*, 530–535.

Roy, B., Frank, M., DeCamp, P., Miller, M., & Roy, D. (2015). Predicting the birth of a spoken word. *Proceedings of the National Academy of Sciences, USA*, *112*, 12663–12668.

Rumelhart, D. E., & McClelland, J. L. (1985). On learning the past tenses of English verbs. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 2, pp. 216–271). Cambridge, MA: MIT Press.

Sachs, J. (1983). Talking about the there and then: The emergence of displaced reference in parent–child discourse. In K. E. Nelson (Ed.), *Children's language* (Vol. 4, pp. 1–28). Mahwah, NJ: Erlbaum.

Sag, I., Wasow, T., & Bender, E. (1999). *Syntactic theory: A formal introduction*. Stanford, CA: CSLI Press.

Suppes, P. (1974). The semantics of children's language. *American Psychologist*, *29*, 103–114.

Tardif, T., Fletcher, P., Liang, W., Zhang, Z., Kaciroti, N., & Marchman, V. A. (2008). Baby's first 10 words. *Developmental Psychology*, *44*, 929–938.

Theakston, A., Lieven, E., Pine, J., & Rowland, C. (2001). The role of performance limitations in the acquisition of verb-argument structure: An alternative account. *Journal of Child Language*, *28*, 127–152.

Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge, England: Cambridge University Press.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Tomasello, M., & Olguin, R. (1993). Twenty-three-month-old children have a grammatical category of noun. *Cognitive Development*, *8*, 451–464.

Valian, V. (1986). Syntactic categories in the speech of young children. *Developmental Psychology*, *2*, 562–579.

Valian, V., Solt, S., & Stewart, J. (2009). Abstract categories or limited-scope formulae? The case of children's determiners. *Journal of Child Language*, *36*, 743–778.

von Humboldt, W. (1999). *On language: On the diversity of human language construction and its influence on the mental development of the human species* (P. Heath, Trans.). Cambridge, England: Cambridge University Press. (Original work published 1836)

Vosoughi, S., & Roy, D. (2012). An automatic child-directed speech detector for the study of child language development. In *Interspeech 2012* (pp. 2478–2481). Retrieved from http://www.isca-speech.org/archive/archive_papers/interspeech_2012/i12_2478.pdf

Yang, C. (2013). Ontogeny and phylogeny of language. *Proceedings of the National Academy of Sciences, USA*, *110*, 6324–6327. doi:10.1073/pnas.1216803110

Zipf, G. (1935). *The psycho-biology of language: An introduction to dynamic philology*. Cambridge, MA: MIT Press.