Quality of early parent input predicts child vocabulary 3 years later

Erica A. Cartmill[a,1], Benjamin F. Armstrong III[b], Lila R. Gleitman[b,1], Susan Goldin-Meadow[a], Tamara N. Medina[c], and John C. Trueswell[b]

[a]Department of Psychology, University of Chicago, Chicago, IL 60637; [b]Department of Psychology, Institute for Research in Cognitive Science, University of Pennsylvania, Philadelphia, PA 19104; and [c]Department of Psychology, Drexel University, Philadelphia, PA 19104

Children vary greatly in the number of words they know when they enter school, a major factor influencing subsequent school and workplace success. This variability is partially explained by the differential quantity of parental speech to preschoolers. However, the contexts in which young learners hear new words are also likely to vary in referential transparency; that is, in how clearly word meaning can be inferred from the immediate extralinguistic context, an aspect of input quality. To examine this aspect, we asked 218 adult participants to guess 50 parents' words from (muted) videos of their interactions with their 14- to 18-mo-old children. We found systematic differences in how easily individual parents' words could be identified purely from this socio-visual context. Differences in this kind of input quality correlated with the size of the children's vocabulary 3 y later, even after controlling for differences in input quantity. Although input quantity differed as a function of socioeconomic status, input quality (as here measured) did not, suggesting that the quality of nonverbal cues to word meaning that parents offer to their children is an individual matter, widely distributed across the population of parents.

language acquisition | word learning | SES

Children's vocabularies vary greatly in size by the time they enter school (1, 2). Because preschool vocabulary is a major predictor of subsequent school success (3), this variability must be taken seriously and its sources understood. Some of this variability resides in the individual capacities and temperament that infants bring to the word learning task (4, 5). However, environmental influences are also bound to play instrumental roles. Accordingly, we examined the contextualized speech input parents provide to infants during the second year of life as a potential source of the massive vocabulary differences found at school entry.

It is already known that the sheer quantity of linguistic input is an important determinant of vocabulary size; overall, the more words children hear early in development, the larger their subsequent vocabularies. This relationship holds true both for types (different words) and tokens (number of words heard, including repetitions) (6, 7). These quantity differences are correlated with socioeconomic status (SES). Children from low SES homes are typically exposed to fewer words early in development (8, 9) and have smaller vocabularies at school entry than children from high SES homes (10).

Taken alone, the correlation of vocabulary size with amount of input is puzzling because as a general rule language learners do not seem to require a large number of exposures to a word to acquire its meaning (11). In experimental settings, for example, children have been shown to acquire and retain a new word heard only once or a very few times (12–14). The likelihood, then, is that certain exposures to a new word are especially informative, supporting secure and rapid inferences to meaning. For example, common sense insists that it will be easier to acquire the meaning of "zebra" in the visual presence of a zebra ("There goes a zebra!") than in its absence ("Let's visit the zebras in the zoo"). To that extent, an environment that maximizes this "here and nowness" of speech—its high quality, or referential transparency—might be expected to boost the rate of early word learning independent of the number of times a child hears each word.

Accordingly, we report here on the influence of such input-quality factors on word learning in infancy, asking: (i) how the referential transparency of input varies across families; (ii) how this variation impacts child vocabulary size 3 y later at school entry; and (iii) how this relationship interacts with SES and the quantity of linguistic input.

The first task in such an inquiry is to pin down a relevant description of input quality. Although, as just acknowledged, having an object in plain view when it is linguistically labeled must be useful, by itself this criterion is far from sufficient, because, in real life, every observed situation is replete with objects, events, properties, and relations. Thus, the learning problem becomes one of selection among many possible interpretations of the speaker's actual referential intention (15, 16).

Linguistic context is one well-known source of information that helps resolve this selection problem in older infants and toddlers who already have a considerable vocabulary and some appreciation of how words are put together syntactically in the language being learned (17–20). However, these potentially informative linguistic cues are largely inaccessible to the rank novices (14- to 18-mo-olds) whose learning environments we study here (21). In contrast, nonverbal cues (e.g., the presence and salience of a word's referent, whether it is being looked at, pointed to, or manipulated by the adult interlocutor) are available from earliest infancy and can, at times, simplify the task of inferring word meaning from situational context (22–25). Attempts have been made to investigate these nonverbal cues using detailed coding systems (26–28). Although such investigations have often been informative (for review, see ref. 29), the subtlety, variety, and sheer number of possible nonverbal cues make it difficult to decide a priori whether any particular contextualized utterance is in principle "good for learning" (26, 27). Rather than attempt to enumerate and classify these nonverbal environmental cues to a word's referent, we estimated the referential transparency of learning instances in infant-directed speech using a reliable and well-validated overall measure of referential transparency—how easily the meaning of a word can be inferred from nonlinguistic context alone—the so-called Human Simulation Paradigm (HSP) (16, 30).

In HSP, adult participants watch muted 40-s video clips ("vignettes") of actual parent-child interactions and try to guess the "mystery word" (indicated by a beep) that the parent uttered at a particular point in the video. The accuracy of these guesses is taken as a measure of how easily the meaning of the word can be inferred from situational context; that is, its quality as a learning

PSYCHOLOGICAL AND COGNITIVE SCIENCES

opportunity. Note that this measure of input quality is independent of the number of words a parent produces and thus is not confounded with input quantity. Prior studies show close agreement in adult and child responses in HSP, suggesting that they are sensitive to the same contextual properties (31, 32).

Evaluating Word-Learning Quality and Its Outcomes

Here we used HSP to test the following general prediction: Families who provide a greater proportion of high-quality word-learning opportunities early in childhood produce better vocabulary outcomes in their children. This prediction was tested in four steps: input sampling, quality assessment, quantity assessment, and language-attainment assessment.

Input Sampling. Fifty parent-child dyads from a stratified SES sample in the Chicago area were video recorded in their homes in whatever situations arose during two 90-min observation sessions at 14 and 18 mo of age (2, 33, 34). For each family, we randomly extracted 10 40-s vignettes, each showing the parent saying a different concrete noun (e.g., dog, ball) (*SI Text*) directed to the child.

Quality Assessment. Potential quality of each vignette was assessed via HSP: that is, showing the muted vignettes to adults ($n = 218$) who tried to guess the target word (indicated by a beep) that the parent uttered. Average accuracy was taken as the measure of the quality of that parent's input; a good learning opportunity occurs when the word's meaning can be readily inferred from its situational environment.

Quantity Assessment. We calculated each parent's average number (in tokens) of spoken words per minute during the two observation sessions at 14–18 mo.

Language-Attainment Assessment. A standardized vocabulary assessment [Peabody Picture Vocabulary Test (PPVT) (35)] was administered to the children at 54 mo. Scores served as the outcome measure at school entry, allowing us to ask how individual differences in quality and quantity of parent input early in development (14–18 mo) correlate with children's comprehension vocabulary 3 y later.

Results and Discussion

We found that the quality of socio-visual input to word meaning varied widely across the 50 parents, with parent HSP accuracy scores ranging from 5–38% (mean 22% ± 8%). Thus, some parents' speech to their offspring rarely contained highly informative contextual cues to meaning, whereas others' did so relatively often. Strikingly, this parent-input quality difference at child age 14–18 mo significantly correlated with the children's vocabularies at 54 mo (linear regression, $r^2 = 0.12$, $P = 0.014$) (Fig. 1*A*). [In all cases of linear regressions reported in this article, corresponding regressions were also computed removing any family identified as an "outlier" (operationalized as having a Cook's distance score greater than 0.15). Unless otherwise noted, these analyses yielded the same patterns of statistical significance.]

Quantity of speech input also varied widely across the 50 parents, with the number of words per minute produced during the 90-min videotaping sessions at ages 14 and 18 mo ranging from 6.3 to 97.0 words per minute (mean 37.8 ± 20.0). Furthermore, replicating previous work (6), this measure of quantity correlated with children's vocabularies at 54 mo (linear regression, $r^2 = 0.13$, $P = 0.011$).

The positive relationship between quality of input and later child vocabulary held even after statistically controlling for effects of quantity on this outcome measure [$r^2 = 0.22$, t(quality) = 2.40, P(quality) = 0.020] (Fig. 1*B*). When taken together (in a multiple regression), the measures of quality and quantity of early parent input at 14–18 mo accounted for 22% of the variance in children's
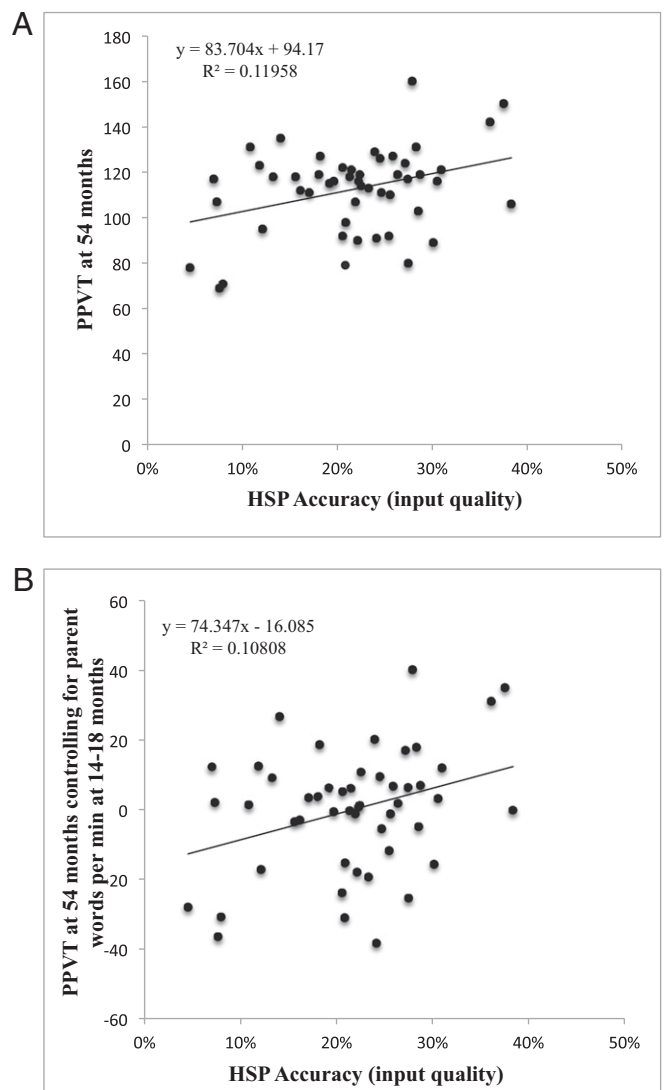
Fig. 1. Effect of quality of early input at 14–18 mo on child comprehension vocabulary at 54 mo. (*A*) Quality of word learning instances (average HSP accuracy per family) at 14–18 mo predicts child comprehension vocabulary (PPVT) at 54 mo. (*B*) This effect holds even after controlling for the quantity of early input (average parent words per minute at 14 and 18 mo). Each point represents a single family ($n = 50$).

vocabularies at 54 mo, a surprisingly strong relation given that 3 y, and presumably many life changes, intervened between assessments.

We next asked how quantity and quality combine to impact linguistic growth. First, quantity of parent input did not correlate with the HSP measure of input quality (Pearson correlation, $r = 0.11$ $P = 0.454$). That is, parents who talked more to their children did not, as a group, provide proportionally more or less high-quality word-learning instances. Second, quality and quantity did not interact with (or moderate) each other when predicting vocabulary outcome in a multiple regression that simultaneously included the main effects of quality, quantity and the interaction term ($P = 0.874$). That is, early quality and quantity accounted for different aspects of the variance found in the later vocabulary outcome measure.

However, parents who talk more are, by definition, offering their children more words, and the more words a child hears, the more likely it will be for that child to hear a particular word in a high-quality learning situation. That is, sheer parental talkativeness increases the likelihood that a child will encounter

high-quality learning instances for particular words during natural parent-child conversational exchanges. To illustrate this effect, we generated an estimate of the total number (as opposed to the proportion) of high-quality learning instances a child receives by multiplying the HSP measure of input quality for each parent by that parent's input quantity measure. This number significantly correlated with child vocabulary at 54 mo (simple linear regression, $r^2 = 0.20$, $P = 0.001$; that is, the interaction term of quality times quantity reliably predicts vocabulary outcome on its own). Numerically, this is a stronger correlation than the correlation between quantity and child vocabulary at 54 mo (0.20 vs. 0.13), suggesting that the number of quality word-learning instances may matter more than the total number of words heard; however, these two correlation coefficients were not significantly different from each other in a comparison test of overlapping coefficients ($z = -1.26$, $P = 0.21$), which takes into account the fact that these two predictors are also highly correlated with each other ($r^2 = 0.72$, $P < 0.001$).

Did quality and quantity of early word-learning opportunities covary with the SES of our families? Primary caregiver education (range: <10–18+ y) and family income (range: <\$7,500 to \$100,000+) were combined to create a single SES variable (see *SI Text* for details). Replicating previous work (10), we found that SES correlated with the quantity of parent input ($r^2 = 0.15$, $P = 0.006$) (Fig. 2A). However, we found that SES did not correlate with our proportional (HSP) measure of quality ($r^2 = 0.005$, $P = 0.625$) (Fig. 2B). When both quality of parent input and family SES were considered as predictors of child vocabulary (in a multiple regression), the model accounted for 35% of the variance in vocabulary at 54 mo, and both quality and SES remained significant predictors [$r^2 = 0.35$, t(quality) = 2.64, P(quality) = 0.011]. Thus, parent SES and our measure of quality are not related to each other, and account for different aspects of the variance found in child vocabulary size at 54 mo.

Input to Naïve Learners. The advantage of studying input to word learning using videos of natural parent-child interactions is that it captures the real variability and complexity of the environments in which word learning naturally occurs. The drawback is that HSP participants have the opportunity to observe the child's reaction following the parent's utterance. If the child already knows the word the parent is saying and reacts accordingly, our HSP participants might be using that reaction to guess the mystery word rather than, or in addition to, using the utterance's concordance with objects and actions in view. For example, in a video where the target word was "nose," a parent might say: "Show me your nose." Vignettes in which the child already knows the word and points to her nose will be easier to guess correctly than videos in which the child does not know the word. Although we made extensive efforts to exclude such interactions from our vignette sample, it is still possible that child's prior knowledge of target words, rather than the informativity of parental use in context, could be driving HSP accuracy scores. If so, then the relationship between parent input quality and later child vocabulary could be nothing more than child vocabulary at 14–18 mo predicting later child vocabulary.

To address this problem, we assessed, on a child-by-child basis, whether there was evidence that the child was already producing any of the randomly sampled test words, using both observation of production (had the child ever spontaneously uttered this word during the 90-min video observation sessions) and a well-validated parental report measure of productive vocabulary [the MacArthur-Bates Communicative Development Inventory (CDI) (36)]. Discarding all vignettes for which there was evidence of prior production reduced both the number of vignettes per family (mean = 6.62 ± 1.78, range 3–10) and the number of families in the dataset ($n = 42$), but resulted in a set of vignettes ($n = 295$) that we could be
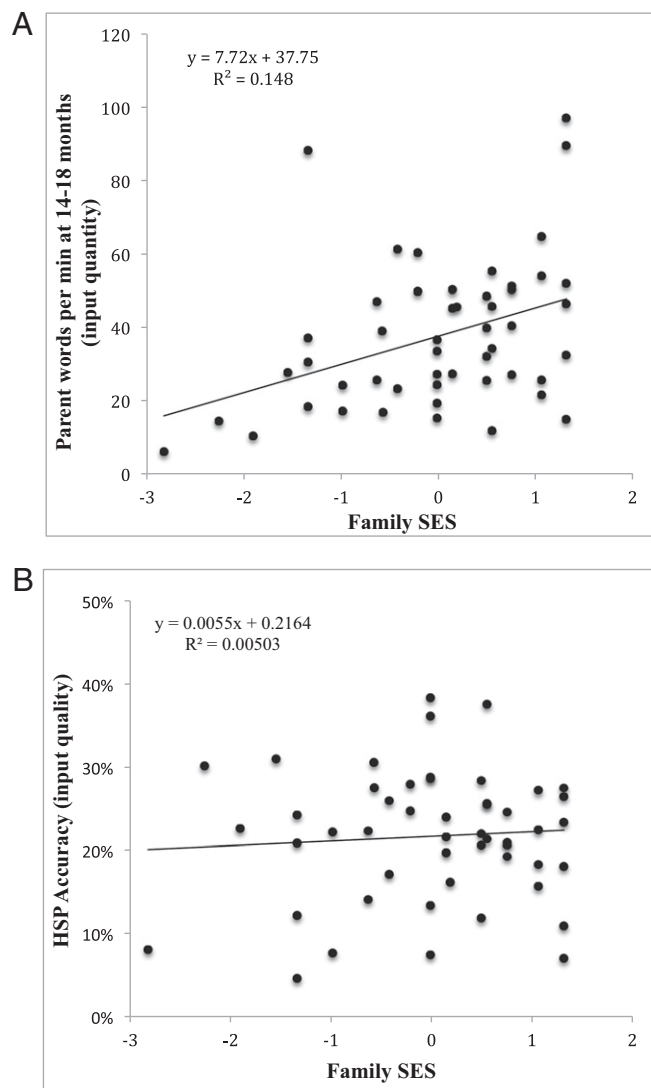
Fig. 2. Effect of SES on quantity and quality of parent input at 14–18 mo. (*A*) SES of family predicts average words per minute uttered by the child's parent during two 90-min recording sessions at 14 and 18 mo. (*B*) SES of family does not predict quality of parent input at 14 and 18 mo (accuracy measure calculated from HSP). Each point represents a single family ($n = 50$).

reasonably certain contained target words not yet in the child's vocabulary.

All analyses were then rerun on this reduced dataset of words not yet known by the child. If our reported correlation between quality and vocabulary outcome were actually a product of known words, we would expect this finding to be eliminated in this subset of unknown words. If, instead, the correlation reflects our ability to capture the quality of potential word-learning instances, the unknown subset should show a correlation similar to the correlation seen in the larger dataset. In all cases, the significant effect of parent input quality remained for unknown words. Specifically, average HSP accuracy for unknown words varied across the 42 parents and resembled the range of scores in the whole dataset (1–35%, mean 19% ± 9%). The effect of parent input quality (HSP accuracy) on children's vocabularies at 54 mo remained (linear regression, $r^2 = 0.11$, $P = 0.036$) (Fig. 3*A*). Conversely, parent input quality derived from videos of known words did not correlate with children's vocabularies (linear regression, $r^2 = 0.03$, $P = 0.306$).
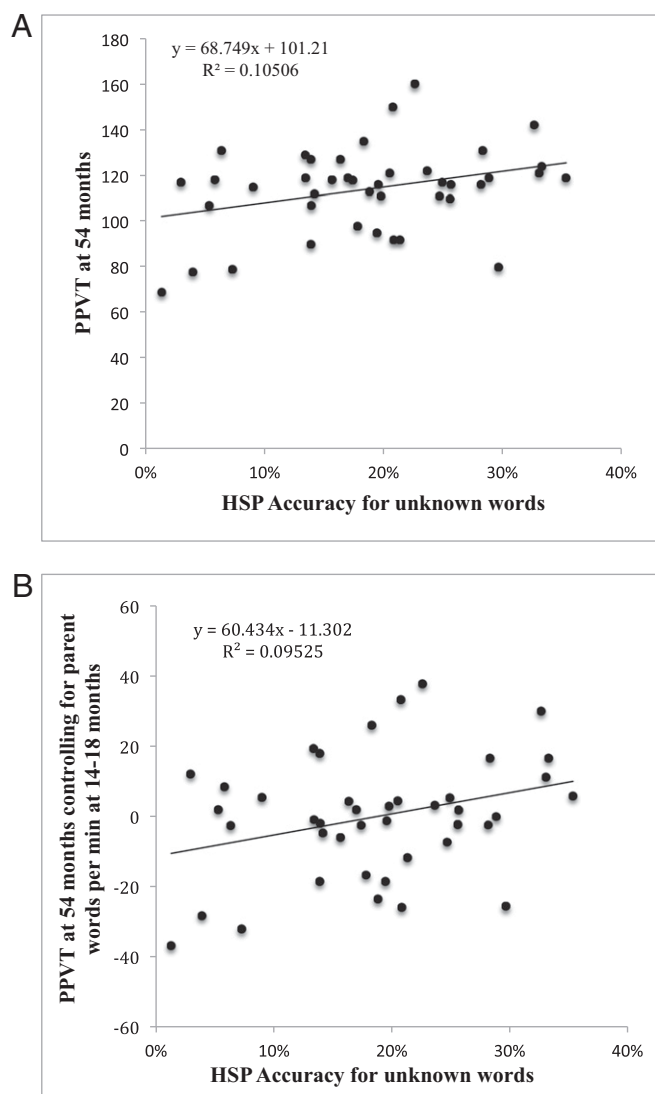
Fig. 3. Effect of quality of parent input for unknown words on child comprehension vocabulary at 54 mo. (*A*) Quality of unknown word learning instances (average HSP accuracy of vignettes in which child does not know the target word) at 14–18 mo predicts child comprehension vocabulary (PPVT) at 54 mo. (*B*) This effect holds even after controlling for the quantity of early input (average parent words per minute at 14 and 18 mo). Each point represents a single family (*n* = 42).

As before, the quality of parent input to naïve learners correlated with later child vocabulary even after controlling for any effects of the quantity of parent input [r^2 = 0.23, *t*(quality) = 2.038, *P*(quality) = 0.048] (Fig. 3*B*). [In a corresponding multiple regression in which four families were removed as numerical outliers, the overall correlation remained significant (r^2 = 0.22, *P* < 0.05), but the quality coefficient was now marginally significant, *t*(quality) = 1.93, *P*(quality) = 0.062.] Also as before, quality of parent input did not correlate with quantity of parent input (Pearson correlation, *r* = 0.10 *P* = 0.520) or with family SES (Pearson correlation, *r* = 0.04 *P* = 0.827).

Thus, the strength of the relation between quality of parent input and later child vocabulary remained when restricting the assessment corpus to those items for which the child participants were less likely to have known the meanings of the words. We also controlled for children's total productive vocabulary at 18 mo (using CDI percentile scores) and found that parent input quality

for unknown words continued to predict vocabulary at 54 mo, but CDI vocabulary score did not [r^2 = 0.14, *t*(quality) = 2.347, *P*(quality) = 0.024), *t*(vocabulary 18 mo) = −0.932, *P*(vocabulary 18 mo) = 0.357]. Together, these findings demonstrate that the effect of quality of parent input on future vocabulary size cannot be explained by differences in children's vocabularies at 14 and 18 mo.

Exploring Quality. An important next step in this line of research will be to identify the behavioral variables that contribute to high-quality input: that is, to identify which properties of the natural physical environment and which behaviors of a parent/speaker promote accurate reading of referential intent. For example, as one would certainly predict, vignettes in which the referent was visible (*n* = 178) at the beep were easier to guess than those in which the object was not visible (*n* = 117, 20% vs. 4% HSP accuracy). Laboratory experiments point to additional, social-pragmatic cues to guessing referential intent, including parent attention to the referent and parent gesture (deictic and nondeictic) (e.g., refs. 22 and 23). Notably, we replicate these findings in our own vignettes involving words not known by the child. In particular, vignettes in which parents were attending to the target object at the beep (*n* = 63) were significantly easier to guess than vignettes in which the object was present but the parent did not attend to it (*n* = 115): 40% vs. 9% HSP accuracy, *t*(176) = −4.59, *P* < 0.001.

Conclusions. Our findings provide significant and unique information about the conditions supporting vocabulary growth early in life. As quantified by a measure of interpretability in context (HSP), the coordination of word use with socio-visual aspects of parent-child interaction is a potent facilitator for the discovery of first word meanings. Grossly speaking, such an effect is self-evident. After all, how could novice word learning happen except by aligning the language sounds with their environmental contingencies? What is less expected is our first finding of great variability in individual parents' natural tendency to provide this contextual support regularly (ranging from 5 to 38% quality: that is, contextually informative, input), with increased informativity having a clear positive effect on vocabulary size by the time formal schooling begins 3 y later. Moreover, this variability in informativity is apparently an individual matter unrelated to SES, and thus seemingly uncorrelated with the more overt teaching styles and picture-book environments that are more prominent in high SES households (37). The positive effect of SES on vocabulary outcome is more likely related to the greater amount of talking by parents to their children in higher SES homes, which, in turn, increases the number of quality learning instances encountered overall.

As an important methodological note, the HSP measure of informativity has primarily been used with adult participant-judges, and these "ideal observer" findings have been assumed to be applicable to the child word-learning case as well. This assumption gains its a priori reasonableness from the finding that the kinds of words adults identify easily in the HSP paradigm are just the kinds of words infants learn early in language development (30). However, some commenters have questioned this inferential link between adult laboratory performance and child learning. The present study, which also used adult participant-judges, goes a long way toward alleviating any such provisos, putting HSP on a new and firmer explanatory footing. We have demonstrated here that adult HSP performance using videotapes of parental input to a sample of 14- to 18-mo-olds significantly predicts these children's vocabulary attainment 3 y later. In other words, the results of the adult HSP predicts word learning in the real world. Situational evidence supporting vocabulary growth seems to work in closely related ways, no matter the sophistication of the observer, and indifferently to whether they are acquiring a first or a second language (38, 39).

We can now return to the apparent paradox described in our introductory remarks: On the one hand, child word learning appears to be amazingly rapid during a period extending from infancy at least throughout late childhood. Indeed, the sheer math of vocabulary acquisition (mean receptive vocabulary of ~12,000 words by the sixth birthday, 23, 40) is understandable, especially if word learning typically requires only one or a very few exposures to words in context for the meaning to be acquired. Recent experimental evidence supports a model [termed "propose but verify" (31, 41)] that comports with these findings by suggesting that learners form a single conjecture about word meaning given a context. Here "context" can be either the immediate situational environment or the immediate linguistic environment for the word [the latter becomes increasingly effective as the child acquires and builds grammatical knowledge (30)]. However, this hypothetical procedure, operating summarily on very little input data, must have a systematic way of avoiding false mappings that may seem plausible in any adventitiously observed situation. The learning procedure succeeds by a species of filtering: it discards or ignores low-quality encounters that happen along the way, thus preventing them from entering into the search for meaning. However, if this fast-mapping procedure is the one that best characterizes child learners, why should the sheer quantity of words children hear correlate so well with vocabulary attainment [as previously reported (6, 7) and as replicated herein]? This frequency-sensitivity finding might suggest a relatively "slow-mapping" learning procedure during which children (indiscriminately) aggregate across both more and less informative encounters with a word, eventually identifying the correct meaning probabilistically on a best-fit basis [as implemented in, for example, simple associationist models (42)]. There is, in fact, suggestive evidence for this kind of cumulative procedure, not only in the effects of frequency that we and others have reported, but also in the generalizations children make on the basis of frequent contexts [e.g., young children often insist that cousins are necessarily children and that uncles are adults (43)].

The present results offer a way to reconcile these positions. As we have documented herein, vocabulary attainment is significantly correlated with quality input, where "quality" is defined as referential transparency (operationalized as HSP score), an effect that is independent of "quantity" (and independent of SES as well). This result comports with the propose-but-verify fast-mapping procedure that makes conjectures only in the presence of high-quality information. However, arguably this procedure could lead to a frequency effect on vocabulary attainment all of the same, simply because frequency probabilistically increases opportunities for encountering highly informative learning instances: the only kind of instance that, when coupled with later confirmatory instances, pushes the child's vocabulary forward.

Summarizing, the present study is consistent with the position that words are learned via a relatively sudden, determinative, and insightful procedure, rather than by brute-force statistical machinery that is accumulating and cross-tabulating all observed instances and utterances across extended periods of time. Still, this fast-mapping machinery, just like statistical-learning machinery, entails an advantage for the more talkative over the more taciturn learning environments within a family and, derivatively, a correlation between SES and vocabulary attainment during the preschool years (because high SES families, as a group, provide higher quantity of input speech). However, the deeper point is that the environment that provides quality input supports efficient learning and that variations in quality, as we have defined and documented it, are observed across the SES spectrum.

In conclusion, the present findings dovetail nicely with laboratory studies of early language learning at the phonetic and word level, showing that an infant's ability to pick up on information in the social context is strongly linked to learning a novel natural language (22–25). The present study shows that this social information—quality encounters in which interpretation is transparent to socio-visual inspection—is delivered to varying degrees by parents in natural settings, and that its presence predicts later language skills.

Materials and Methods

Parent-Child Pairs. Fifty parent-child pairs were included in the analyses. All parents gave written informed consent to participate in the longitudinal study and for their videos to be used in additional research. Children were typically developing (27 males, 23 females). The pairs were from a larger sample of 63 families participating in a longitudinal study of language development (33, 34); see *SI Text* for details. Families were chosen to be representative of the ethnicity and income ranges of the greater Chicago area. All children were raised monolingual English speakers. Families were visited in their homes every 4 mo from child age 14 to 58 mo, and were video recorded for 90 min at each visit. During visits, families engaged in their normal daily activities, ranging from book reading and puzzle play to meals and bathing. Selection of families for present analyses ($n = 50$) was based on the following exclusion criteria (*SI Text*): parents did not permit videos to be used in future research ($n = 6$); the parent was not a native-English speaker ($n = 1$); the child did not take the PPVT test at 54 mo ($n = 6$).

The following measures were taken of each family (*SI Text*): (*i*) SES as determined by a principle component analysis (33) combining the education level of the primary caregiver (range: 10–18 y; mean = 15.96, SD = 2.16) and family annual income (range: under $7,500 to over $100,000; mean = $64,000, SD = $30,000); (*ii*) Quantity of parent's linguistic input was the average number of words per minute in child-directed speech produced by the parent at the 14- and 18-mo visits; (*iii*) Child vocabulary outcome was measured using the PPVT at 54 mo (35). We used the PPVT scores from 54 mo because they reflect child vocabulary just before school entry.

Vignette Selection for HSP. A total of 560 40-s muted videos (vignettes) served as target stimuli in the HSP study below. These vignettes came exclusively from the 14- and 18-mo-old visits. Each vignette was an example of a parent uttering one of the 40 most common concrete nouns in the transcript sample, uttered usually within a sentence context (e.g., *Can you give me the book?*). Vignettes were aligned so that ~30 s into the video, the parent uttered the target word (at which point a beep was inserted). If the parent uttered the target word more than once during the 40-s vignette, each instance of the target word was replaced by a beep. We considered this local repetition of target words to be a feature of the quality of the context and thus allowed it to vary naturally. Previous studies found this duration to be sufficient to understand the gist at the moment the target word was uttered (31). The *SI Text* describes in detail our vignette selection criteria, but in brief, 10 vignettes were selected from each of 56 participating families. (This sample included the six families later excluded for lack of a PPVT test score at 54 mo.) Five filler videos were also selected from each family: that is, 280 fillers. Filler words consisted of verbs, adjectives, quantifiers, or nouns that were not easily visualized, so as to prevent participants from only guessing concrete nouns.

HSP Experimental Design. Because no single HSP participant could reasonably view and respond to all 840 40-s vignettes (9.33 h of video in total), vignettes were split into 15 experimental lists such that each list had no more than one vignette from each family and had no more than four examples of the same word.

HSP Participants. Participants ($n = 218$) were randomly assigned one of the 15 lists consisting of 56 vignettes (including both target and filler words). Participants were undergraduate students (145 female, 73 male) enrolled at the University of Pennsylvania ($n = 159$) or La Salle University ($n = 59$) in Philadelphia. All were native English speakers, coming from relatively diverse SES backgrounds (see *SI Text* for details). The protocol was approved by the Institutional Review Boards at the University of Chicago, the University of Pennsylvania, and La Salle University. Participants provided informed consent and received course credit or payment for participating.

HSP Procedure. After viewing a vignette, participants guessed the "mystery" word for that vignette before viewing the next. Participants were tested individually or in groups, ranging from one to six people. Video was projected on a wall or screen and participants recorded their guesses on paper. See *SI Text* for details.

Analyses. Participant guesses were scored as correct if they were identical to the target word. Abbreviations and plurals were also counted as correct (e.g., phone or phones for telephone), but words that altered the meaning

of the root word were not (see *SI Text* for details, and ref. 31). Participants' responses were used to calculate the guessing accuracy (average number of correct guesses) for each video clip. The guessing accuracies of a parent's 10 target videos were then averaged to create an average HSP accuracy for that parent. This average was used as the measure of quality of input for each family.

We used linear regressions to examine the relationships between: (*i*) the quantity of parent input, (*ii*) the quality of parent input, (*iii*) child vocabulary at 54 mo, and (*iv*) family SES. In particular, we analyzed the ability of input quality and quantity to predict child vocabulary at 54 mo.

For analyses involving only those target words not yet known by the child, child knowledge of HSP target words was determined by combining the words the child actually produced during the two 90-min observation sessions with the words parents reported the child produced at 14 and 18 mo. Recorded production was measured during the 90-min video observation sessions at 14 and 18 mo (i.e., a child spontaneously produced a target word during that time). Parent report of production was determined through the

MacArthur-Bates CDI (36), in which parents marked any words they believed their child had said. The CDI was given to parents at every observation session, but not all parents returned the CDI at each session. Only vignettes of parents who completed the CDI at 18 mo were included in the unknown word dataset. If either observation or parent report indicated that the child produced an HSP target word, it was deemed known. These "known-word vignettes" were excluded from analyses where our focus was on parent input to naïve learners.

1. Fenson L, et al. (1994) Variability in early communicative development. *Monogr Soc Res Child Dev* 59(5):1–173, discussion 174–185.
2. Rowe ML, Raudenbush SW, Goldin-Meadow S (2012) The pace of vocabulary growth helps predict later vocabulary skill. *Child Dev* 83(2):508–525.
3. Anderson RC, Freebody P (1981) Vocabulary knowledge. *Reading Comprehension and Education*, ed Guthrie JT (International Reading Association, Newark, DE), pp 77–117.
4. Stromswold K (2001) The heritability of language: A review and metaanalysis of twin, adoption and linkage studies. *Language* 77(4):647–723.
5. Tsao FM, Liu HM, Kuhl PK (2004) Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Dev* 75(4):1067–1084.
6. Huttenlocher J, Haight W, Bryk A, Seltzer M, Lyons T (1991) Early vocabulary growth: Relation to language input and gender. *Dev Psychol* 27(2):236–248.
7. Hurtado N, Marchman VA, Fernald A (2008) Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Dev Sci* 11(6):F31–F39.
8. Hart B, Risley TR (1995) *Meaningful Differences in the Everyday Experiences of Young American Children* (Brookes, Baltimore, MD).
9. Hoff-Ginsberg E (1991) Mother-child conversation in different social classes and communicative settings. *Child Dev* 62(4):782–796.
10. Hoff E (2003) The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Dev* 74(5):1368–1378.
11. Carey S (1978) The child as word learner. *Linguistic Theory and Psychological Reality*, eds Halle M, Bresnan J, Miller GA (MIT Press, Cambridge, MA), pp 264–293.
12. Spiegel C, Halberda J (2011) Rapid fast-mapping abilities in 2-year-olds. *J Exp Child Psychol* 109(1):132–140.
13. Nappa R, Wessel A, McEldoon KL, Gleitman LR, Trueswell JC (2009) Use of speaker's gaze and syntax in verb learning. *Lang Learn Dev* 5(4):203–234.
14. Heibeck TH, Markman EM (1987) Word learning in children: An examination of fast mapping. *Child Dev* 58(4):1021–1034.
15. Fodor JA (1981) The present status of the innateness controversy. *Representations*, ed Fodor JA (MIT Press, Cambridge, MA), pp 257–316.
16. Gleitman LR (1990) The structural sources of word meaning. *Lang Acquis* 1(1):3–55.
17. Naigles L (1990) Children use syntax to learn verb meanings. *J Child Lang* 17(2):357–374.
18. Gertner Y, Fisher C, Eisengart J (2006) Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychol Sci* 17(8):684–691.
19. Gleitman LR, Cassidy K, Nappa R, Papafragou A, Trueswell JC (2005) Hard Words. *Lang Learn Dev* 1(1):23–64.
20. Hoff E, Naigles L (2002) How children use input to acquire a lexicon. *Child Dev* 73(2):418–433.
21. Bruner J (1974) From communication to language—A psychological perspective. *Cognition* 3(3):255–287.
22. Baldwin DA (1991) Infants' contribution to the achievement of joint reference. *Child Dev* 62(5):875–890.
23. Bloom P (2000) *How Children Learn the Meanings of Words* (MIT Press, Cambridge, MA).
24. Kuhl PK (2007) Is speech learning 'gated' by the social brain? *Dev Sci* 10(1):110–120.
25. Brooks R, Meltzoff AN (2008) Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: A longitudinal, growth curve modeling study. *J Child Lang* 35(1):207–220.
26. Allen SEM (2000) A discourse-pragmatic explanation for argument representation in child Inuktitut. *Linguistics* 38(3):483–521.
27. Frank MC, Tenenbaum JB, Fernald A (2013) Social and discourse contributions to the determination of reference in cross-situational word learning. *Lang Learn Dev* 9(1):1–24.
28. Laakso ML, Poikkeus AM, Katajamäki J, Lyytinen P (1999) Early intentional communication as a predictor of language development in young toddlers. *First Lang* 19(56):207–231.
29. Hoff E (2006) How social contexts support and shape language development. *Dev Rev* 26(1):55.
30. Gillette J, Gleitman H, Gleitman L, Lederer A (1999) Human simulations of vocabulary learning. *Cognition* 73(2):135–176.
31. Medina TN, Snedeker J, Trueswell JC, Gleitman LR (2011) How words can and cannot be learned by observation. *Proc Natl Acad Sci USA* 108(22):9014–9019.
32. Piccin TB, Waxman SR (2007) Why nouns trump verbs in word learning: New evidence from children and adults in the Human Simulation Paradigm. *Lang Learn Dev* 3(4):295–323.
33. Rowe ML, Goldin-Meadow S (2009) Differences in early gesture explain SES disparities in child vocabulary size at school entry. *Science* 323(5916):951–953.
34. Rowe ML, Levine SC, Fisher JA, Goldin-Meadow S (2009) Does linguistic input play the same role in language learning for children with and without early brain injury? *Dev Psychol* 45(1):90–102.
35. Dunn LM, Dunn LM (1997) *Peabody Picture Vocabulary Test* (American Guidance Service, Circle Pines, MN), 3rd Ed.
36. Fenson L, et al. (1993) *The MacArthur Communicative Development Inventories: User's Guide and Technical Manual* (Brookes, Baltimore, MD).
37. Fletcher KL, Reese E (2005) Picture book reading with young children: A conceptual framework. *Dev Rev* 25(1):64–103.
38. Snedeker J, Geren J, Shafto CL (2007) Starting over: International adoption as a natural experiment in language development. *Psychol Sci* 18(1):79–87.
39. Johnson JS, Newport EL (1989) Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognit Psychol* 21(1):60–99.
40. Anglin JM (1993) Vocabulary development: A morphological analysis. *Monogr Soc Res Child Dev* 58(10, Serial No. 238):1–186.
41. Trueswell JC, Medina TN, Hafri A, Gleitman LR (2013) Propose but verify: Fast mapping meets cross-situational word learning. *Cognit Psychol* 66(1):126–156.
42. Yu C, Smith LB (2007) Rapid word learning under uncertainty via cross-situational statistics. *Psychol Sci* 18(5):414–420.
43. Keil FC (1989) *Concepts, Kinds and Cognitive Development* (Bradford Books, Cambridge, MA).