Corpus Evidence for Age Effects on Priming in Child Language

Jeffrey Gerard (jeffreygerard@gmail.com)
Frank Keller (keller@inf.ed.ac.uk)
School of Informatics, University of Edinburgh
10 Crichton Street, Edinburgh EH8 9AB, UK

Themis Palpanas (themis@disi.unitn.eu)
Department of Information Engineering and Computer Science
University of Trento, Via Sommarive 14, I-38123 Povo, Italy

Abstract

Structural priming, the tendency to repeat previously uttered syntactic structures, can give insight into human language processing and acquisition. We report two corpus-based studies of children's structural priming that test the following claim of the item-based account of language acquisition: as older children generalize over structures, priming increases with age. A hypothesis derived from this claim, viz., that the lexical boost effect decreases with age, is also tested. We fit mixed-effects logistic regression models on data from children aged 2 to 7.5 years from the CHILDES corpus. We demonstrate structural priming of arbitrary syntactic structures for the first time in child language data. We also find evidence that priming increases with age, but fail to confirm the hypothesis that the lexical boost effect decreases with age.

Keywords: Syntactic priming; Child language; Corpus studies; Mixed models; Age effects in language acquisition.

Introduction

Priming occurs when an initial stimulus, called the *prime*, causes a bias towards a related stimulus later on. Adaptation to the prime manifests itself in the latter stimulus, the *target*, which is comprehended faster or more accurately, or produced more frequently. A wide range of priming effects has been documented, including the priming of words, syntactic structures, and discourse patterns. The phenomenon is neither intentional nor conscious (Bock & Loebell, 1990). Establishing which aspects of a linguistic stimulus adapt to priming—and which ones do not—gives insight into the mental representation of language and the process by which speakers comprehend and produce sentences.

The vast majority of priming research has been carried out with adults, but there are some recent studies that investigate priming in children (e.g., Savage, Lieven, Theakston, & Tomasello, 2003; Huttenlocher, Vasilyeva, & Shimpi, 2004; Kemp, Lieven, & Tomasello, 2005). Such studies make it possible to examine the development of linguistic representations, based on how priming effects change over the time course of language acquisition, i.e., with the age of the child. Priming can therefore be used as a tool to test specific questions about human language acquisition.

One of the key questions in language acquisition is whether grammatical rules are acquired conjointly with individual words or, alternatively, syntactic knowledge is abstract from lexical knowledge. In the latter case, the question arises of the source of knowledge of the abstract structure of a language, since children's only input to language acquisition is the lexical expressions that they hear. Tomasello's (2000) *item-based hypothesis* proposes that children's early language consists of word-for-word chunks copied from adults' phrases, from which they only gradually abstract patterns and therefore grammar rules. An alternative view is that all children are born with a *universal grammar* (Chomsky, 1980); this theory suggests that abstract grammatical knowledge is innate in the human brain, and merely needs to be parametrized during the course of language acquisition.

This paper explores the item-based hypothesis by studying structural priming in corpora of child language. If a child adapts to structural priming—that is, the child shows a tendency to reuse syntactic constructions heard or produced recently—then this indicates that the child is using old syntactic representations to express new ideas with different words. The item-based hypothesis predicts that this behavior should increase with the age of the child: if syntactic development is a gradual shift from lexically dominant phrase repetition towards generalized grammatical rules, then structural priming should be more frequent in older children, who have more abstract syntactic representations available.

In a well-studied phenomenon called *lexical boost*, structural repetition rises when the target and the prime share a content word, i.e., lexical adaptation boosts structural adaptation (Pickering & Ferreira, 2008). We hypothesize that if grammatical abstraction is thought of as curtailing reliance on words, then priming may show decreased effects of lexical boost as children age.

In this paper, we test both hypotheses: that overall priming increases with age, and that the lexical boost effect decreases with age.

Background

Many experimental studies create an atypical context of language use, requiring the participant to respond to a number of similar trials, where the high repetition of trials may condition participants to become more practiced in their responses, or alternatively, participants may show fatigue. Priming studies, in particular, often present made-up nonce words and observe participants' comprehension or use of them (e.g., Brooks & Tomasello, 1999; Kemp et al., 2005). Teaching a participant a novel word requires multiple exposures which means multiple primes, and it is not clear what effect additive priming might have. Likewise, several priming experiments, especially with children (e.g., Savage et al., 2003; Kemp et al., 2005; Huttenlocher et al., 2004), entail both hearing and then repeating every prime, again double-priming all targets. Corpus studies are not subject to these confounds, and they can help verify that a phenomenon observed in a few children in a few contexts can be generalized to child language as a whole.

With few exceptions, experimental and corpus studies alike have looked for priming of a small set of specific syntactic alternations—different syntactic forms that express the same semantics—providing very limited coverage of grammar. Recent corpus studies have overcome this limitation, and have found that priming is a more general phenomenon (Reitter, Moore, & Keller, 2006; Reitter, 2008), and that less frequent structures show more priming that more frequent ones (the inverse frequency effect).

In the current paper, we present the first corpus-based investigation of priming in children. In the first of two studies, we replicate an experimental study of the priming of passive and active constructions in children (Savage et al., 2003). Our second study generalizes these results by modeling adaptation to the priming of arbitrary structures. The studies bear on the item-based hypothesis of language acquisition. In particular, we investigate the role of a child's age as a predictor of priming, and consider the influence of lexical similarity.

Modeling Methodology

We used mixed-effects logistic regression to model how various explanatory variables affect structural repetition between pairs of sentences from the CHILDES corpus.

Data

The CHILDES corpus (MacWhinney, 2000) contains over 100 databases of transcriptions of face-to-face interactions between young children and their caretakers. The corpus studies described in this paper used a subset of these databases that contain multiple interviews with a child over different dates, so that priming could be compared at different ages of the same child.[1] For naturalness, the phrase "the corpus" will refer to this subset of CHILDES. The corpus comprises utterances from 84 child speakers, as well as speech from their adult interlocutors.

The most current collection of CHILDES transcripts as of April 17, 2009 was processed to remove structures containing unrecognized words, babble, test words, and fillers (onomatopoeia and child-invented word forms that could be recognized were kept). Certain types of clitics were separated to correspond with morphosyntactic annotations (e.g., *they'll* ⇒ *they will*), as were assimilations (e.g., *wanta* ⇒ *want to*). Disfluencies, retracings, and repetitions were kept.

CHILDES includes annotations of morphemes and syntactic categories, which are automatically generated by supervised taggers (MacWhinney, 2000). This is in turn used to generate labeled dependency structures based on grammatical relations between words (Sagae, Lavie, & MacWhinney, 2005). Sagae et al. evaluate the dependency hierarchy accuracy to be 90.1% on child language transcripts.

Mixed-Effects Logistic Regression

We used mixed-effects logistic regression to identify which variables influence priming in our corpus. Our dependent variable Y is a binary variable that indicates whether there is structural repetition between two sentences ($Y = \text{TRUE}$) or not ($Y = \text{FALSE}$). Logistic regression is a generalization of linear regression that predicts the logit of the probability p that Y is TRUE, as a function of explanatory variables $X_1 \ldots X_N$:

$$\text{logit}(p) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_N X_N.$$

The logit link function is $\text{logit}(p) = \ln\left(\frac{p}{1-p}\right)$. Its inverse is the logistic function, ensuring that as a probability $0 \leq p \leq 1$:

$$\begin{aligned} p &= \text{logit}^{-1}\left(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_N X_N\right) \\ &= \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_N X_N)}} \end{aligned}$$

Mixed-effects regression allows the optional inclusion of *random effects* in order to generalize beyond the groups represented in a limited statistical sample. Modeling random effects allows for the possibility that, say, different children have different repetition behaviors, or that each child's repetition behavior changes in different ways as he or she gets older. In the first example, a random effect variable CHILD would be defined to model trends that are specific to each one of its grouping factors: Abe, Abigail, Adam, Allison, etc.

For each possible value g of the random-effects grouping variable, let α_g be the deviation of the dependent variable's mean for group g from the entire dataset's mean; α_g is a random variable from a normal distribution with a mean of zero and unknown variance. α_g is added to each datum taken from group g, effectively adjusting the intercept of regression independently for each group so that uncontrolled effects specific to each group do not unfairly distort the overall model.

Model Specification and Fitting

In the corpus studies reported here, we fitted mixed models with random intercepts grouped by the child speaker of target utterances, which were further nested by database to account for random effects of different annotators, interview strategies, etc. In Study 2, random intercepts were also defined for the syntactic structure being investigated in each datum, which is particularly important because the frequencies of the structures vary greatly, approximating a Zipf distribution.

The corpus data is strongly biased towards younger children, with relatively few utterances from children above five years old. Unlike natural cases of sparsity (e.g., spoken language uses far fewer passive sentences than active ones) the sparsity of data for older children is an artifact of CHILDES. Still, it presents a potential problem, as the model-fitting algorithms had to deal with higher variance for older children.

[1] The subset comprised the following databases: Bloom73, Brown, Demetras1, Demetras2, Feldman, Gathercole Gleason, Kuczaj, MacWhinney, Sachs, Suppes, Wells. The Wells database contains British English; all the others contain American English.

Models were fit using Laplace's method by the `lme4` software package for the R programming environment. All explanatory variables were centered around the mean to reduce multicollinearity between higher-order interactions and their constituent main effects. We built minimal models by dropping non-significant explanatory variables (unless explicitly relevant to the experiment, or necessary as the component of a significant higher-order interaction).

Study 1: Priming of Active and Passive Voice

To confirm that mixed-model regression analysis of corpus data can provide an insight into structural priming similar to what can be accomplished in experiments, we replicated an experimental design utilized by Savage et al. (2003) and compared the qualitative results of the two methods.

The experiment of Savage et al. (2003, experiment 1) proceeded as follows: In interviews with 84 children from age 2;11 to 7;1 (*years*; *months*), children heard and repeated a prime sentence—either active or passive—describing some transitive action depicted in a cartoon. Then they were shown another cartoon of a different action with different participants and asked "What's happening?". The target sentences the children produced in response were classified as PASSIVE or ACTIVE. Experimenters also varied the amount of lexical overlap that the child could potentially find between the given prime sentence and the child-produced target.

Method

All sentences in the specified subset of the CHILDES corpus (see Modeling Methodology above) were automatically identified as active, passive, or other, guided by heuristics. Whereas all passives primed in Savage et al.'s (2003) experiment included an agentive *by*-clause (e.g., *The ball got caught* ***by the net***), the corpus contains only four examples of children using a passive form with expressed agent, one of which is recitation from a storybook. Agentive *by*-clauses are optional in English, and their sparsity appears to be representative of natural language production (Huttenlocher et al., 2004). Accordingly, the present study considered agentless passives (e.g., *I got caught*) along with agentive passives.

The Savage et al. experiment considered "only the first sentence-like utterance … produced after exposure to each prime sentence," so we also compared only adjacent utterances from the corpus. Only pairs where the target was spoken by a child from age 2;0 to 7;6 were included; the potential primes were spoken by adults and children of all ages, but were always spoken in the presence of the target child. Furthermore, pairs were omitted from the analysis if either of the two sentences contained a negation or was a *wh*-question, which were not used by Savage et al., or if a sentence was not identified as obviously passive or active. A contingency table of the remaining pairings already makes clear that an active prime is much more likely than a passive one to precede an active target; see Table 1.

To answer the main questions of whether children's priming is dependent on their age and on lexical overlap, we fit a mixed-effects logistic regression model where the voice of the target sentence was predicted by the following main effects and their interactions:

- PRIME: the voice of the prime utterance (ACTIVE or PASSIVE);
- AGE: the child's age represented as decimal years with precision to the day;[2]
- LEXBOOST: the ratio of the number of words in common between both utterances to the total number of words in the target utterance;
- PRIMETYPE: CP for comprehension-production priming (another speaker produces the prime and the child comprehends it and produces the target) or PP for production-production priming (the child produces both the prime and the target).

Table 1: Frequencies of adjacent prime-target pairings in Study 1, where the target was spoken by a child.

		Target	
		Active	Passive
Prime	Active	359	13
	Passive	14	13

Table 2: Study 1 parameter estimates. Explanatory variables estimate the logit of the probability that TARGET (the utterance immediately following PRIME) is passive.

	β	$p(> \lvert z \rvert)$	
(Intercept)	−4.447	$\ll 0.001$	***
PRIME[PASSIVE]	1.597	0.082	·
AGE	0.351	0.179	
LEXBOOST	−1.373	0.274	
PRIME[PASSIVE]:LEXBOOST	16.285	< 0.002	**

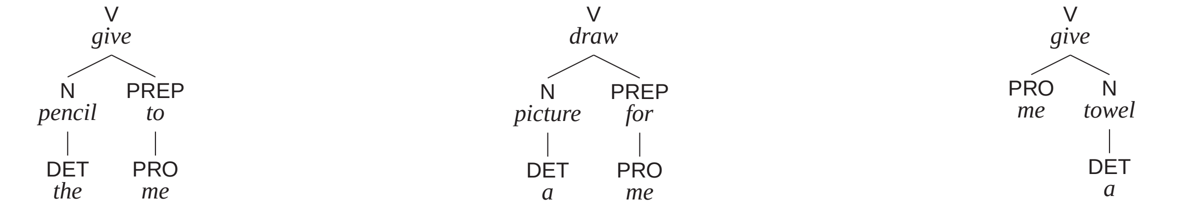
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, · $p < 0.1$

Results

Table 2 above gives the coefficients of the mixed model together with significance values. We find a significant interaction of PRIME and LEXBOOST. All other interactions and the main effect of PRIMETYPE were evaluated and found to be non-significant regressors, so the model was refit without them. In particular, the model shows no influence of a child's age on his production of active or passive sentences, with or without active or passive primes (no main effect of AGE, no interaction PRIME:AGE).

The dependent variable TARGET was mapped such that passive targets yield TRUE and active targets yield FALSE. The positive coefficient for the PRIME[PASSIVE]:LEXBOOST interaction therefore means that together, a passive prime and lex-

[2] Where a child's age was specified to only monthly precision, the median value of 15 days after the start of that month was assumed.



Give the pencil to me. (a) Prepositional dative

(Do you want to) draw a picture for me? (b) Prepositional dative

Give me a towel. (c) Ditransitive dative

Figure 1: Two prepositional datives share the same structural analysis, which differs from a ditransitive dative

ical boost *increase* the probability that the target will be realized as a passive (as opposed to an active). The marginally significant main effect of PRIME[PASSIVE] lends weak evidence for priming of passives also in the absence of lexical overlap.

Discussion

Our method of modeling the corpus is not identical to the analyses Savage et al. (2003) performed on their experimental data, but the results are comparable in qualitative terms. Savage et al. performed separate analyses of variance (ANOVA, an instance of linear regression), one for each target voice. For both voices, Savage et al. found reliable main effects of PRIME. The reliable interaction between PRIME and LEXBOOST we found in the corpus data was also present in Savage et al.'s ANOVA for passive targets. Meanwhile, PRIME and LEXBOOST formed part of a significant three-way interaction with AGE only in their ANOVA of active targets; they further broke down the active targets to find the PRIME:LEXBOOST interaction in their three- and four-year-old participants but not in six-year-olds. As mentioned above the CHILDES corpus is sparse in data over age five, which likely explains why we did not find any interaction with age.

That the effect of age was only found in active targets suggests that the sparsity of passive targets in both datasets is also important. This weakness cannot be overcome with the studies structured as they are—neither in experimental data nor in corpus data—simply because of the natural sparsity of passives in children's spoken language. Therefore, instead of relying on only a single alternation for insight into children's language, in the next study we investigated children's to priming of arbitrary syntactic structures.

Study 2: Priming of Arbitrary Structures with Decay

Most priming studies to date have only considered structures for which a semantically equivalent alternation exists. This limits the generality of conclusions that can be drawn, and data sparseness is a potential problem, as illustrated above. In the present study, we therefore use an approach that does not require the existence of an alternation, asking instead whether the appearance of a prime structure increases the probability that the *same structure* will appear again.

We define priming in probabilistic terms: the appearance of a prime structure increases the conditional probability that the same structure will appear again:

$$p\left(S_{\text{prime}} \mid S_{\text{target}}\right) > p\left(S_{\text{prime}}\right)$$

where $p\left(S_{\text{u}}\right)$ is the prior probability that an arbitrary structure S will appear in any utterance u . Using this approach, general structural priming—not only for specific structures—can be quantified in a single model.

Besides the sparsity of passives, both Study 1 and the experiment on which it was based had another limitation. By considering only adjacent utterances, they treated priming as an immediate phenomenon and ignored its well-documented temporal decay (Branigan, Pickering, & Cleland, 1999; Pickering & Ferreira, 2008). The formalism just presented is easily extended to estimate $p\left(S_{\text{prime}} \mid S_{\text{target}}, d\right)$, that is, the probability that structure S appeared in the d-th utterance before TARGET was spoken (Reitter et al., 2006; Reitter, 2008).

Structural Overlap

To measure repetition of arbitrary syntactic structures, we need a way to identify whether two constructions share the same structure or are syntactically distinct. We used the hierarchical structure supplied in the form of CHILDES's dependency annotation for this purpose, based on evidence that priming relies on shared hierarchical syntactic rules (Bock & Loebell, 1990; Reitter, 2008). However, priming is not sensitive to thematic roles (Bock & Loebell, 1990), so the relation labels in the annotation are not useful. We therefore used the part-of-speech tags from the CHILDES morphological annotation instead, imposed upon the dependency hierarchy. This combination gives the same analysis to those structures typically considered correspondent in priming studies (Figures 1a and 1b) while producing different analyses for their characteristic alternations (Figure 1c vs. 1a and 1b).

For this study, we used the subset of such structures that have exactly three levels. Of this subset, those with very low frequency—fewer than about twenty occurrences over the entire corpus, according to a manual evaluation—were usually incorrect analyses derived from inaccurate annotations in CHILDES (either in the morphosyntactic or the dependency structure). Thus data points corresponding to structures with frequency less than twenty were discarded. This leaves 4,279 unique structures for consideration, representing 81.3% of the original data. No outliers on the high end of the frequency spectrum were discarded, as they were correct analyses.

Method

Each structure S in some child's (age 2;0 to 7;6) utterance t was considered a potential target of adaptation, primed by the structures in all utterances p within the window of the fifteen utterances preceding t. For each combination of t, $S \in t$, and d ($1 \leq d \leq 15$), a record was created of whether S was in $p = t - d$. That is, the model's binary dependent variable represents *repetition* of a certain structure across a certain distance. Consequently, the parameters estimated by the regression model are effects on mere structural repetition. Priming is identified in this formulation by its decay, so only interactions with the variable DIST (which represents d) can be interpreted in terms of priming; specifically, negative coefficients of DIST indicate priming.

Because measuring grammatical abstraction requires differentiating between lexical and structural repetition, data points showing structural repetition resulting from complete lexical repetition (i.e., not differing by at least one word)—one-half a percent of the dataset—were dropped for this study. Structures in the first fifteen utterances of any interview session also were not considered as targets because they may have been influenced by primes not captured in the corpus. The remaining data points were segregated into strata, one stratum for each three-month period of each child. Two-thousand five-hundred data points were randomly sampled from each stratum, unless the stratum contained fewer than 2,500 points, in which case the entire stratum was used.

A mixed-effects logistic regression model was built to correlate structural repetition across distance (DIST) with explanatory variables AGE and PRIMETYPE as described in Study 1 and with the following variables:

- LEXBOOST: a binary variable that is TRUE if the heads of both hierarchical structures use the same *root morpheme* (lemma);
- ln(FREQ): the logarithmically transformed frequency of the structure in the entire corpus.

This experimental setup crucially relies on the assumption that priming decays. Figure 2 plots the sampled probability that an arbitrary structure is repeated between two utterances separated by a variable distance. It clearly shows the probability of repetition diminishes as a function of distance, with higher repetition across shorter distances—in short, structural priming decays.

There is evidence that both human attention and priming decay logarithmically (McKone, 1995). This is supported by Figure 2, and indeed the mixed-effects model yields a better fit when variable DIST is transformed logarithmically than when it is linear.

Results

Table 3 shows parameter estimates of the full model specification. We find a significant, negative coefficient for ln(DIST), showing the decay of priming of arbitrary structures in children's speech. In line with previous research (Reitter, 2008), the significant interaction ln(DIST):ln(FREQ) demonstrates

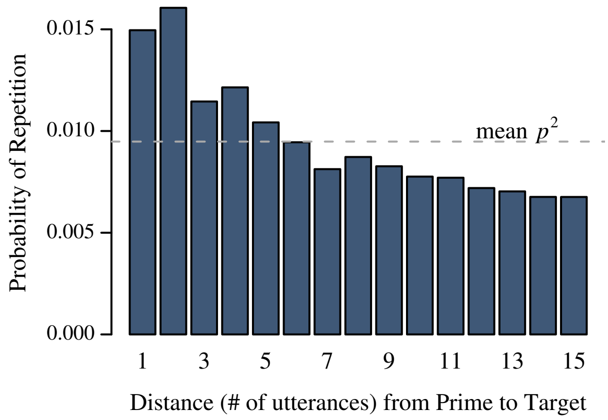


Figure 2: Decay of priming in children. The dashed line is the mean squared prior probability of structures, estimated over comparisons of up to fifteen utterances; since priming raises the probability of a structure above its prior, the true mean will be closer to the repetition probability at DIST = 15.

that less frequent structures show stronger adaptation. Note that the interaction's positive coefficient needs to be interpreted in the context of the negative slope of ln(DIST) (repetition decreases with increasing distance) and the positive slope of ln(FREQ) (less repetition for more frequent structures).

We also observe a significant positive coefficient of ln(DIST):LEXBOOST, suggesting that priming (that is, the decay of DIST) may weaken under lexical boost, all other factors held fixed. Meanwhile, ln(DIST):ln(FREQ):LEXBOOST has a negative coefficient: the decay effect increases with lexical boost and increasing frequency. This means we find the lexical boost effect to be stronger for high-frequency items.

The ln(DIST):AGE interaction is marginally significant ($p = 0.075$), providing only weak evidence for the claim that structural priming increases with age. More convincing support of this prediction is offered by the significant ln(DIST):ln(FREQ):AGE interaction. Its positive coefficient means that priming (the decay of DIST) becomes stronger if age increases while frequency decreases, or weaker as age and frequency increase. In other words, the inverse-frequency effect is stronger for older children than for younger children.

ln(DIST):AGE:LEXBOOST is not significant. We therefore find no evidence for our suggestion that lexical boost may influence structural priming differently as children gradually abstract grammar from phrasal repetition.

The marginally significant ln(DIST):PRIMETYPE interaction hints that children may be more inclined to repeat their own previous constructions (PRIMETYPE = PP) than primes by another speaker.

Discussion

This model shows that priming of arbitrary structures is evident in children, a population in which priming of only a few syntactic alternations had been studied previously. This study also provides an estimate that priming's main efficacy lasts

Table 3: Study 2 parameter estimates. Explanatory variables estimate the logit of the probability of structural repetition. **ln(DIST)** terms are emphasized to remind that this model provides insight to priming only through the DIST variable.

	β	$p(>\lvert z \rvert)$	
(Intercept)	−6.299	$\ll 0.001$	***
ln(DIST)	−0.423	$\ll 0.001$	***
ln(FREQ)	0.701	$\ll 0.001$	***
AGE	−0.191	$\ll 0.001$	***
LEXBOOST	2.738	$\ll 0.001$	***
PRIMETYPE[PP]	0.487	$\ll 0.001$	***
ln(DIST):ln(FREQ)	0.107	$\ll 0.001$	***
ln(DIST):AGE	−0.056	0.075	·
ln(DIST):LEXBOOST	0.114	0.049	*
ln(DIST):PRIMETYPE[PP]	−0.075	0.056	·
ln(FREQ):AGE	0.030	0.007	**
ln(FREQ):LEXBOOST	−0.230	$\ll 0.001$	***
ln(FREQ):PRIMETYPE[PP]	−0.150	$\ll 0.001$	***
AGE:LEXBOOST	0.046	0.349	
AGE:PRIMETYPE[PP]	−0.031	0.579	
LEXBOOST:PRIMETYPE[PP]	0.171	0.017	*
ln(DIST):ln(FREQ):AGE	0.022	0.026	*
ln(DIST):ln(FREQ):LEXBOOST	−0.056	0.009	**
ln(DIST):AGE:LEXBOOST	−0.035	0.344	
ln(FREQ):AGE:LEXBOOST	−0.057	< 0.001	***
ln(FREQ):AGE:PRIMETYPE[PP]	0.062	< 0.001	***
AGE:LEXBOOST:PRIMETYPE[PP]	−0.186	0.003	**

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, · $p < 0.1$

around six utterances, during which it shows strong decay and after which its decay is negligible (see Figure 2). Most importantly, this study enables us to quantify priming effects according to age during first language acquisition.

Crucially, the results do not support the conjecture offered in the Introduction that structural priming's reliance on lexical boost decreases as children age. It is important to bear in mind that this conjecture is not strictly predicted by the item-based hypothesis, which does not specify precisely what types of analogies children must make to abstract a grammar from word patterns. Kemp et al. (2005) provide evidence similar to our results, observing in one experiment that two-year-olds adapted to structural priming without regard to lexical influence.

On the other hand, we did find evidence that overall structural priming increases with age. If this is true, it supports the item-based hypothesis of language acquisition which holds that over time children gradually abstract grammatical rules from the sentences they hear.

Conclusion

This paper reported two corpus-based studies of structural priming during first language acquisition. Study 1 replicated an experiment on passive/active priming in children, and found similar effects in corpus data to those reported experimentally (Savage et al., 2003). Both studies tested the hypothesis that structural adaptation increases with age. Study 2 found evidence for this claim, though the change is not as large as might be expected by an item-based account of language acquisition. Neither study supports our conjecture, influenced by the item-based hypothesis, that the lexical boost effect should decrease with age, as children move from lexicalized to abstract syntactic knowledge.

References

Bock, K., & Loebell, H. (1990). Framing sentences. *Cognition*, *35*(1), 1–39.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (1999). Syntactic priming in written production: Evidence for rapid decay. *Psychonomic Bulletin and Review*, *6*(4), 635–640.

Brooks, P. J., & Tomasello, M. (1999). Young children learn to produce passives with nonce verbs. *Developmental Psychology*, *35*, 29–44.

Chomsky, N. (1980). *Rules and representations*. New York: Columbia University Press.

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

Kemp, N., Lieven, E., & Tomasello, M. (2005, June). Young children's knowledge of the "determiner" and "adjective" categories. *Journal of Speech, Language, and Hearing Research*, *48*(3), 592–609.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd Edition ed., Vol. 2: The Database). Mahwah, NJ: Lawrence Erlbaum Associates.

McKone, E. (1995). Short-term implicit memory for words and non-words. *Journal of Experimental Psychology: Learning Memory and Cognition*, *21*(5), 1108–1125.

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, *134*(3), 427–459.

Reitter, D. (2008). *Context effects in language production: Models of syntactic priming in dialogue corpora*. Unpublished doctoral dissertation, School of Informatics, University of Edinburgh.

Reitter, D., Moore, J. D., & Keller, F. (2006). Priming of syntactic rules in task-oriented dialogue and spontaneous conversation. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 685–690). Vancouver.

Sagae, K., Lavie, A., & MacWhinney, B. (2005). Automatic measurement of syntactic development in child language. In *Proceedings of the 42nd Meeting of the ACL* (pp. 197–204). Association for Computational Linguistics.

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2003, November). Testing the abstractness of children's linguistic representations: lexical and structural priming of syntactic constructions in young children. *Developmental Science*, *6*(5), 557–567.

Tomasello, M. (2000). The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences*, *4*(4), 156–163.