Developmental Change in the Relationship Between Lexical and Grammatical Acquisition

Mika Braginsky
mikabr@stanford.edu
Department of Psychology
Stanford University

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

Virginia Marchman
marchman@stanford.edu
Department of Psychology
Stanford Uiversity

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

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aenean volutpat eu dui et bibendum. Curabitur ut neque vel nulla dignissim pellentesque. Sed lacinia mi sit amet libero pellentesque cursus. Sed ut vulputate augue. Quisque porta leo pulvinar, congue massa at, dapibus metus. Maecenas malesuada eleifend metus quis sagittis. Vivamus fermentum rhoncus sodales. Aliquam vitae pharetra lacus. Duis quis leo id orci porttitor elementum. Integer dictum ex enim, sed convallis tellus ultricies sed. Vestibulum pretium convallis fermentum. Sed ac sapien et magna dignissim aliquet et sed turpis.

Keywords: foo; bar; baz

Introduction

[Intro material written by Mike:] Does abstract structure in language emerge from the interaction of individual words, or are syntactic structures represented separately? On lexicalist theories of grammatical development, syntactic structure emerges from graded generalizations on the basis of lexical items and there may be little or no representation of syntactic rules or regularities per se, at least early in development (Tomasello, 2000; 2003). Even if syntactic structures are eventually represented, these representations should be directly related to their support in more concrete lexical structure (Bannard et al., 2009; Bod, 2010??). In contrast, on theories like principles and parameters, grammar is predicted to emerge independently from lexical knowledge and on its own (largely maturational) timetable. On these theories, older children should have more syntactic competence, largely independent from the amount of language input they receive and hence from the size of their vocabulary.

Developmental data can help resolve this conflict by providing estimates of the relationship between grammar, lexicon, and age. Indeed, in a now-classic set of analyses, Bates & Goodman (1995) showed a systematic relationship between ...

[Intro material written by Mika:] [Sentence about how all grammar generalizations must involve both input and internal stuff]. However, theories of grammatical development differ in the role that they assign to experience and internal processes and the emphasis that they place on one or the other. One one hand, in the more nativist and domain-specific tradition [stuff about generative grammar and Principles and Parameters (cite Atoms of Language and Chomsky); basically everything is maturation]. On the other hand, [stuff about emergentism and lexicalist grammar (cite Bates and Goodman 1999 and maybe some Tomasello)].

[Stuff about how those theories make different predictions and how those predictions can be explored in empirical studies of lexicon-grammar relationship. If there's age-related change, functions should be different at different age points. If there's no age-related change, functions should be the same.]

[What functions? How two relate lexicon to grammatical development? Two different metrics: 1) usage of grammatical forms and constructions; 2) composition of lexicon by grammatical function.]

[How to measure any of these things? CDI of course! Justification for CDI (Fenson, 1994), justification for complexity section (Dale, 1991; Dale et al, 1989) and word form section (???).]

[Previous work on lexicon-grammar: well-established tight connection between lexicon and grammar, same non-linear function everywhere in English (Bates et al, 1994; Fenson et al, 1994, Italian (Caselli, Casadio, Bates, 2000), Hebrew (Maital, Dromi, Sagi, Bornstein, 2000), Icelandic (Thordardottir, Weismer, Evans, 2002), Spanish (Jackson-Maldonado et al, 2003)]. [Bates and Goodman (1999) conclude that grammar is entirely predictable from lexicon. But we find additional effect of age, suggesting there's a tight link but there's also something developmental going on.]

[Previous work on lexical composition: English (Bates et al, 1994), other languages (???).]

General Methods

CDI Form Database

:

[Talk about wordbank, all the sources of our data, give some numbers of kids.]

[Talk in more detail about the CDI and the specifics of the languages and forms.]

Data Sources

CDI Measures

In general CDI forms contain both vocabulary checklists and other questions relevant to the child's ...

Developmental Interactions in Syntax and Morphology

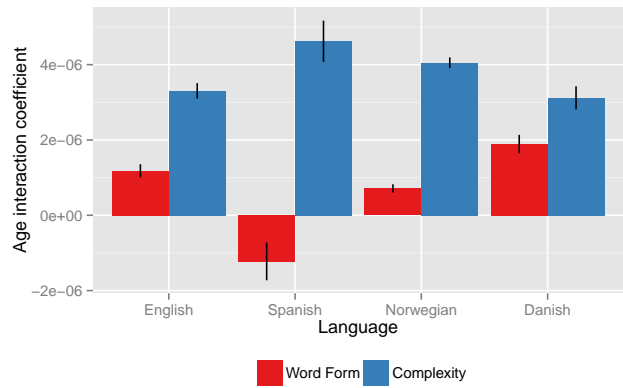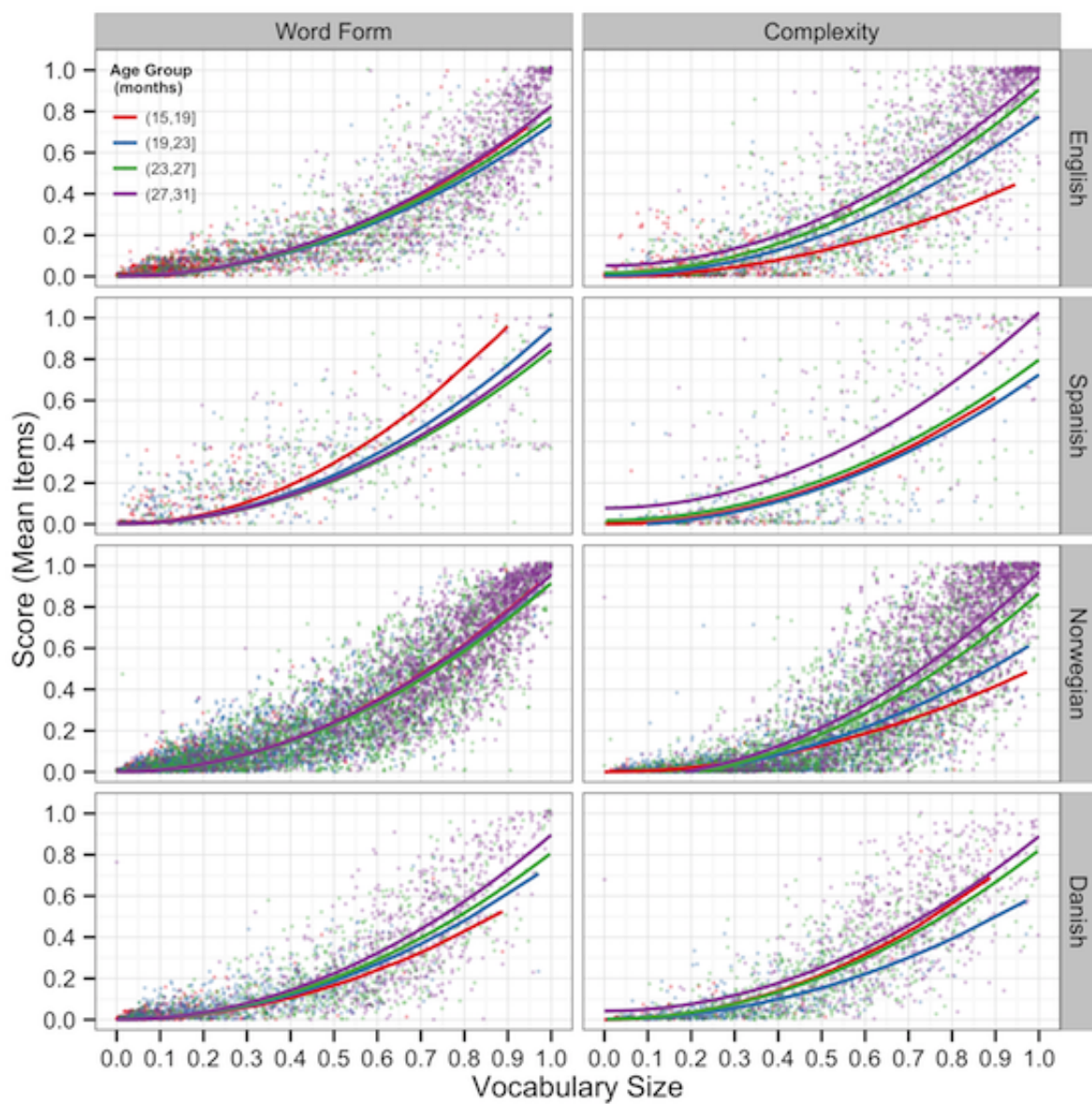[Introduce and scaffold results.]

Figure 2: Coefficients

Figure 1: Grammar!

Developmental Interactions in Vocabulary Composition

Early vocabulary development is typically characterized by learning of names for caregivers and common objects, while l ater in development, children tend to increase their vocabulary by increasing the proportion of predicates (verbs and adjectives). This over-representation of nouns has been found across a number of analyses and in a variety of languages, though not all (Caselli et al., 1995).[1] For our purposes we are interested in using these analyses of vocabulary composition to test for the same kind of age-related differences that we found in the complexity and word-form analyses reported above.

We predict that the proportion of verbs in children's vocabulary should be relatively more affected by age than nouns. Concrete nouns are hypothesized to be learned initially from both co-occurrences between words (Yu & Smith, 2007) and by social cues to reference to particular objects (Bloom, 2002). On neither of these accounts should syntactic information be a primary information source (though of course syntax might be more informative for abstract nouns). In contrast, for verbs, syntax has been argued to be primary in meaningl learning. On the syntactic bootstrapping hypothesis (Gleitman, 1990; ?, ?), verbs are learned by mapping the syntactic structure of utterances to the thematic structure of observed events, for example by noticing that the subject of a sentence matches the agent in one particular ongoing event but not another ("the cat is fleeing the dog" matches FLEES(CAT, DOG) but not CHASES(DOG,CAT)). Thus, if syntactic development is related in some way to age, we should see larger age effects on verb representation than noun representation.

Analysis

Each CDI form contains a mixture of words in different classes. We adopt the categorization of Bates et al. (1994), who split words in nouns, predicates (adjectives, adverbs, and verbs), function words, and other words. Then for each child's vocabulary, we compute the proportion of the total words in each of these categories that they are reported to produce. This analysis yields a set of proportions for each child.

Again following Bates et al. (1994), for each of the four languages in our sample, we plot these proportions against total vocabulary. These functions are shown in Figure 3: each dot represents an individual child's knowledge of a particular class, while curves show the relationship between that class and the entirety of the vocabulary. If categories grow independently of one another, these curves should approximate the diagonal. This pattern is not what either we or Bates

and colleagues observe however: Across the languages in our sample, nouns are systematically over-represented in smaller vocabularies (shown by a curve that is above the diagonal), while function words—and to some extent, predicates—are under-represented.

Next, we measure the contribution of age to vocabulary composition. We fit a linear model to all children's data for each word class, predicting word-class proportion as a linear and quadratic (FIXME) function of total vocabulary. We then investigated the interaction between total vocabulary and age. Because of our theoretical interest in the relationship between age and syntactic development, we focus here specifically on nouns and verbs. Figure **??** shows coefficients for each of these models across languages. In each model, the age-related interaction coefficient was substantially larger for verbs than for nouns. This asymmetry can be interpreted as evidence that for two vocabulary-matched children, the older would tend to have relatively more verbs than the younger, and this effect was larger for children with overall larger vocabularies.

Discussion

We replicated previous analyses showing an over-representation of nouns in the developing lexicon and a relative under-representation of verbs. We additionally predicted that—if syntactic generalization was in some way tied to age—verbs would show relatively more influence than nouns. This prediction was confirmed across all four languages we examined. Thus, this analysis provides additional circumstantial evidence for the relationship between syntactic development and age, independent of the growth of the lexicon.

---

[1]Differences in early vocabulary composition have been argued to emerge from typological differences (e.g., word order, subject drop), and from cultural practices (e.g., focus on picture book reading) (Tardif, Gelman, & Xu, 1999; Gopnik, Choi, & Baumberger, 1996; Choi & Gopnik, 1995)—we are agnostic as to the source of this variability.
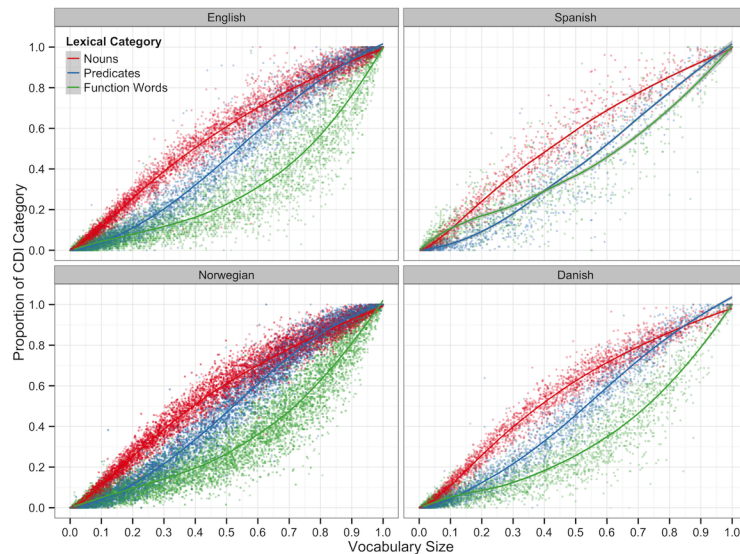
Figure 3: Proportion of a particular CDI category, plotted by total vocabulary size. Each point shows an individual child, with color showing their noun, predicate, and function word vocabulary. Panels show different languages, and curves are smoothing functions using loess.

General Discussion

Acknowledgments

References

References

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., … Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of child language*, *21*(01), 85–123.

Bloom, P. (2002). *How children learn the meanings of words*. Cambridge, MA: MIT Press.

Caselli, M., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L., & Weir, J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development*, *10*(2), 159–199. doi: 10.1016/0885-2014(95)90008-x

Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in korean: A cross-linguistic study. *Journal of child language*, *22*(03), 497–529.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, 3–55.

Gopnik, A., Choi, S., & Baumberger, T. (1996). Cross-linguistic differences in early semantic and cognitive development. *Cognitive Development*, *11*(2), 197–225.

Tardif, T., Gelman, S. A., & Xu, F. (1999). Putting the noun bias in context: A comparison of english and mandarin. *Child Development*, *70*(3), 620–635.

Yu, C., & Smith, L. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, *18*(5), 414–420.