New Frontiers in Computational Models of Grammatical Development

Micah B. Goldwater[1] (micahbg@gmail.com), Scott Friedman[2], Dedre Gentner[1], Ken Forbus[2]
Department of Psychology[1]; Department of Electrical Engineering & Computer Science[2], Northwestern University
Evanston, IL 60208 USA

Cynthia L. Fisher[1] (clfishe@cyrus.psych.illinois.edu), Michael Connor[2], Dan Roth[2]
Department of Psychology; Department of Computer Science, University of Illinois at Urbana-Champagne
Champaign, IL 61820 USA

Franklin Chang (Franklin.Chang@liverpool.ac.uk)
School of Psychology, University of Liverpool
Liverpool, L69 7ZA UK

Gary S. Dell (gdell@cyrus.psych.illinois.edu)
Department of Psychology, University of Illinois at Urbana-Champagne
Champaign, IL 61820 USA

Keywords: computational models; language development; syntax; thematic roles

Introduction

How children acquire the grammar of their native language has been a central topic in cognitive science since its outset, and has been the focus of much debate. One view assumed an innate Universal Grammar which genetically endowed the child with highly structured knowledge of language (Chomsky, 1965). An opposing position argued against both the assumptions of innate knowledge and structured representations, instead using connectionist architectures with distributed representations to learn grammatical patterns (e.g., Rumelhart & McClelland, 1986).

The field has progressed. There have been many years of rigorous empirical work, detailing the developmental pattern in children. In parallel, AI and cognitive science have made many advances in sophisticated learning algorithms. This symposium brings together models on the forefront of such empirical and computational research. Each model has roots in both sides of the early debate, positing (at least some) structured representations and specifying learning mechanisms. However, the models differ in many crucial ways. They use different computational architectures, learning algorithms, and differ in the knowledge built into the system. These differences in the models reflect and build on different current theories of grammatical development. The models focus on simulating empirical phenomena critical in distinguishing such theories. This symposium presents a unique opportunity to compare these new approaches and invite an open discussion.

Symposium Structure

This symposium will present three computational models of grammatical development. The first talk, by Cindy Fisher, presents a model rooted in *early abstraction* theories of language development (e.g., Fisher, 2002). The model is implemented in a machine learning architecture that uses its innate biases linking syntax and semantics and learned grammatical categories to semantically parse sentences and guide word learning. The second talk, by Franklin Chang will describe a connectionist model that does not require explicit thematic roles or innate linking rules to learn syntax, but instead can acquire language from visual-spatial input. The third talk, by Micah Goldwater, presents a third approach—a usage-based model that uses structured symbolic representations and learns abstract thematic roles via analogical abstraction

The symposium begins with a brief introduction, followed by three presentations of computational models, and concludes with a discussion exploring the issues.

Dedre Gentner will introduce the symposium. She is the Alice Gabrielle Twight Professor of Psychology and Education at Northwestern University.

Gary S. Dell will serve as the discussant. He is Professor of Psychology and Linguistics at University of Illinois at Urbana-Champaign.

We now summarize each talk in turn.

The Origin of Syntactic Bootstrapping: A Computational Model

Syntactic bootstrapping proposes that children use knowledge of sentence structure in sentence interpretation and verb learning. We present a computational model of the origins of syntactic bootstrapping, based on systems for automatic semantic-role labeling (SRL). SRL models learn to identify sentence constituents that fill semantic roles, and to determine their roles, such as agent, patient, or goal. The present 'BabySRL' instantiates the structure-mapping account of syntactic bootstrapping (Fisher et al., 2010). We assume a structure-mapping process between the nouns in a sentence and the core semantic arguments of the verbs, in which children are biased to create one-to-one mappings. Given this one-to-one mapping bias, the number of nouns in the sentence becomes intrinsically meaningful to toddlers. Second, this account proposes that children's representations of sentences, though partially specified, are couched in

abstract terms, permitting generalization of new syntactic learning to new verbs. We used the BabySRL to investigate the consequences of these assumptions for learning from natural corpora of child-directed speech. The results yield strong evidence that partial sentence representations grounded in a set of nouns are useful as a foundation for further learning. We show (1) that such representations support new learning about English word order (e.g., the first of two nouns is typically an agent), and (2) permit children to learn which words are verbs by tracking their argument-taking behavior in sentences.

Cynthia L. Fisher is Professor of Psychology and Linguistics at University of Illinois at Urbana-Champaign.

The Dual-Path Model

A theory of language should unite acquisition, production, and comprehension. To link acquisition to production, Chang (2002) proposed a connectionist model called the Dual-path model. The model used the same learning mechanism for acquiring abstract English syntactic representations to explain structural priming in adult production (Chang, et al., 2006). The model could learn English and Japanese to similar levels (Chang, 2009). Importantly, the model's incremental planning mechanism could also account for the different direction of heavy NP shift in each language. That is, English speakers tend to postpone long noun phrases (NPs) to the end of sentences, while Japanese speakers tend to place these NPs earlier. Finally, a new version of the model has been developed where thematic roles in the message are replaced with arbitrary spatial pointers. The production or comprehension of words depends on the activation of these pointers, and hence they can simulate eye-tracking in scenes. The model can account for different types of anticipation in English and Japanese during eye-tracking in the visual world (e.g., Kamide et al., 2003). The model highlights the role of non-linguistic processes such as implicit learning and spatial binding in language acquisition and use.

Franklin Chang is Lecturer of Psychology at the University of Liverpool.

An Analogical Learning Model of the Development of Thematic Roles & Structural Priming

This model explores the hypothesis that analogical learning processes can account for the abstraction of thematic roles (Goldwater et al., 2011). We use SME (Falkenhainer et al., 1989) and SAGE (Kuehne et al., 2000) to model analogical mapping and generalization. Learning proceeds via incremental comparison of specific examples to reveal their common structure and create generalizations. This allows the model to evolve abstract roles from verb-specific ones. We assess the abstractness of the model's semantic roles by simulating structural priming in sentence production. To construct an utterance for a new event, it first uses analogical retrieval to find utterances with similar semantic structure in memory. It then creates an analogical mapping between events and transfers the previous sentence structure to construct the new utterance. Hence, structural priming is shown when a new utterance has analogous structure to a prime utterance.

We assume a usage-based learning trajectory in which children's initial semantic representations are verb-specific (e.g., Tomasello, 2003). Because shared semantic structure is necessary to show priming, early in training priming occurs only across sentences that share verbs. As the model learns, across-verb priming occurs. For example, initially the model can align *giving* events only with other *giving* events, because the roles do not match those of other verbs. By gradual re-representation and generalization, the model develops a hierarchy of more abstract roles.

Micah B. Goldwater is a postdoctoral fellow in the Department of Psychology at Northwestern University.

References

Chang, F. (2002) Symbolically speaking: A connectionist model of sentence production. *Cognitive Science, 26*, 609-651

Chang, F. (2009) Learning to order words: A connectionist model of heavy NP shift and accessibility effects in Japanese and English. *Journal of Memory and Language, 61,* 374-397

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113*, 2, 234-272

Chomsky, N. (1965), *Aspects of the Theory of Syntax*, Cambridge, MA: MIT Press

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence, 41,* 1-63

Fisher, C. (2002). The role of abstract syntactic knowledge in language acquisition: A reply to Tomasello (2000). *Cognition, 82*, 259-278.

Fisher, C., Gertner, Y., Scott, R., & Yuan, S. (2010). Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science, 1,* 143-149.

Goldwater, M. B., Friedman, S. E., Gentner, D. G., Forbus, K. F., Taylor, J. L M. (2011). An analogical learning model of the development of thematic roles & structural priming. *Presented at the 36th Annual Boston University Conference on Language Development.*

Kamide, Y., Altmann, G.T.M., & Haywood, S. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye-movements. *Journal of Memory and Language. 49*, 133-159.

Kuehne, S. E., Gentner, D. & Forbus, K. D. (2000). Modeling infant learning via symbolic structural alignment. *Proceedings of the Twenty-second Annual Conference of the Cognitive Science Society*, 286-291

Rumelhart, D., & McClelland, J., (1986). On Learning the Past Tenses of English Verbs. in McClelland and Rumelhart (eds). *Parallel Distributed Processing*, 216–271

Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press