Using a Computational Model of Language Acquisition to Address Questions in Linguistic Inquiry

Jane Hill
Department of Computer Science
Smith College, Northampton, MA 01063

In this paper we are presenting a case study to illustrate the way in which our computational model of English language acquisition and language performance in very young children may be used to address questions in linguistic inquiry. Ours is a model which depends on the interaction of language and cognitive domains, and which seeks commonalities between the processes employed in language perception and sensory perception in other modalities, as well as between language production and response to stimuli in other modalities.

1. Computational Models and the Innatist-Empiricist Controversy

Since Noam Chomsky published Syntactic Structures in 1957 the dialogue between the innatists and the empiricists has fundamentally influenced the course of research in language acquisition. The debate between Chomsky and Piaget (Piatelli-Palmarini 1980) served to illustrate how firmly established are the opposing points of view, and how little movement toward reconciliation such debates as these seem to foster. Our model of language acquisition in the two-year-old (Hill 1982, 1983) has some bearing on the above controversy. The model represents a first attempt at building a vehicle which may be employed to examine and experiment with different hypotheses about language learning. Although it is obvious that computational models in and of themselves will not solve any controversies, nevertheless we believe that ultimately models such as ours will have a profound effect on the way in which questions will be posed and the answers that will be found. Even our simple model has already succeeded in stimulating its own healthy controversy.

2. Brief Overview of Model I

We will now describe our very simple model of the acquisition of English in the two-year-old and discuss some questions and some answers suggested by the use of the computational model in analyzing the Adam data which was gathered by Roger Brown, Ursula Belugi, Colin Fraser, and Courtney Cazden prior to 1973 (Brown 1973).

We can only give a very brief overview of the model and its assumptions in this paper. Readers interested in details of the model should consult Hill (1982, 1983) for full particulars of the model including examples of computer output together with corresponding linguistic data collected from a two-year-old child. The psychological validity of the model is defended in Hill and Arbib (1984) and in Hill (1984). The model is described as a member of the class of schema-theoretic models in Arbib, Conklin, and Hill (to appear). The model as described here is fully implemented. The development of the model, however, is an ongoing process. Figure 1 provides a diagram of the components of the model.

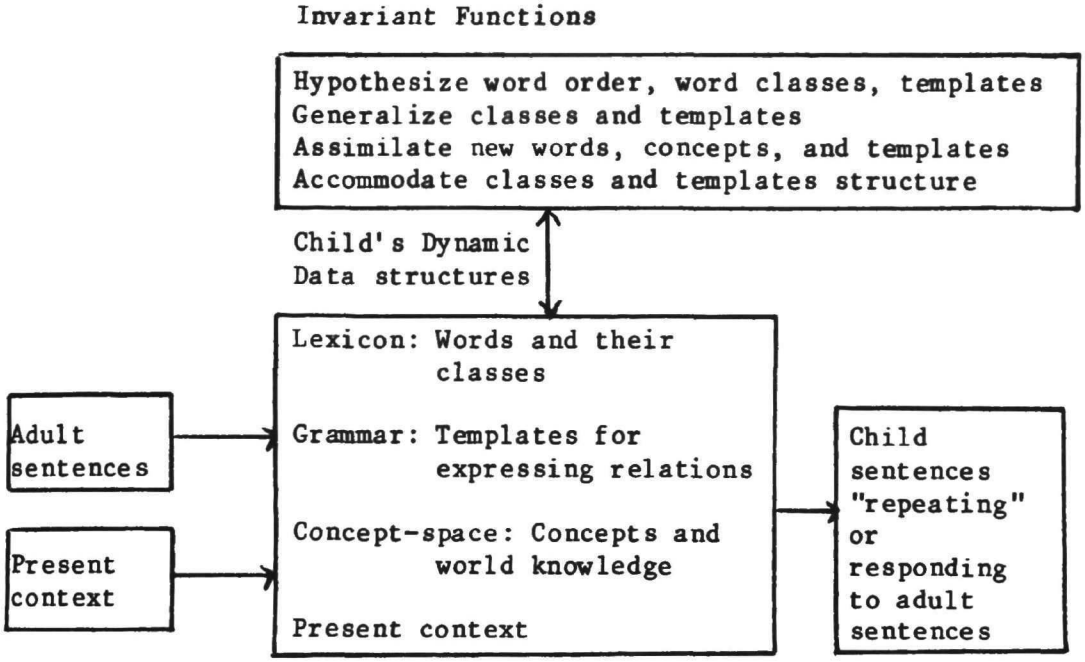

Figure 1. Components of the model

The model takes as its input adult sentences together with indications provided by the user, where relevant, of the physical context in which the sentences are uttered. Output from the model is a representation of child-like sentences repeating or responding to the adult input in accordance with the current state of the model's grammar. The child's knowledge is represented by dynamic data structures encoding the child's lexicon, the child's grammar, the conceptual knowledge of the child, and the physical context of the dialogue. No assumptions were made about the ultimate form of the adult grammar nor about what must be built-in to the model, but a precise account was kept of the knowledge and processes which were found to be necessary to be built-in to the model even for this elementary level of speech. The model notices and employs relations, word order, rules for concatenating relations and deleting words, and as the model grows word classes are formed. An important constraint on the model is that it does not depend on knowledge of adult word classes such as a division between actions and objects. The model requires schemas for word classification and template classification in order to grow, but the actual classes remain flexible and are inferred from the child's language behavior. The processes named above cause successive reorganizations of the grammar and the lexicon. The model suggests one way in which language based initially on cognitive knowledge can grow into a syntactic system which will be increasingly independent of its semantic and cognitive foundations. It is important to note that although the rules embodied in the model are simple, their interaction is complex enough to necessitate the use of a computer model.

The model in its original form (which we shall refer to as Model I) attained only the level of a two-year-old producing sentences of up to six words in length. The grammar of Model I (which we shall refer to as Grammar I) was entirely flat and was made up of rules for forming two-word utterances expressing relations and for combining those two-word relations into utterances which were two to six words in length. The grammar and lexicon

contained no articles and no conjunctions. The natural next question to be addressed by the model was how might the child progress from the flat grammar to a hierarchical one? As a means of examining this question we decided to explore the child's use of conjunctions.

3. Expanding Model I to Accommodate Coordinate Structure

We have examined the data collected by Roger Brown and his colleagues from the child Adam, from the age two years three months up to two years eleven months with particular attention to the use of the conjunction "and" by the adults in the transcribed sessions and by the child Adam. We used as input to our model the body of adult sentences which occurred in the transcribed sessions and which contained the conjunction "and." The output obtained from the model without any additional machinery quite readily matched the Adam data up to the time that Adam himself began to use conjunctions. What follows is a case study of the use of the model in exploring the child's understanding and formation of coordinate structure. The questions which linguists have addressed in the analysis of coordinate structure will be very briefly summarized. For a more complete discussion see Stockwell, Schacter, and Partee (1973).

1. One may employ conjunction reduction whereby redundant elements in a fully sentential coordinated sentence are deleted. For example

    John and Bill left

is derived from

    John left and Bill left

with the first occurrence of the word "left" deleted.

2. One may conjoin phrases and use the conjoined phrase as the subject of the verb "left."
3. One may employ some combination of these tactics.

In their article "The Development of Sentence Coordination", Tager-Flusberg, deVilliers & Hakuta (1982) argue that children's earliest phrasal conjunctions are not plausibly derived by conjunction reduction but are generated by directly combining like constituents by phrase structure rules. It will turn out that the model supports this argument but it should be noted that the claim is by no means universally accepted by linguists. (See for example Lust 1977, Lust and Mervis 1980.)

Our approach to building a model is to address the question of what must be added step by step to the model in order to parallel the developing language of the child. No one would imagine that the simple rules of Model I would suffice for any but the most elementary stage of language. The question we ask of the model is not if it will fail, but rather in what fashion it will fail, and most importantly can the model evolve in such a fashion that the more mature processes evolve from and grow out of the earlier processes. In the transcribed data, Adam at first frequently failed to respond at all to conjoined adult sentences. When Adam began to repeat conjoined utterances he repeated them by omitting "and." All these exchanges were developed by the model without any additional modification to the model's Grammar I or to the lexicon or concept-space of Model I. This was as it should have been.

At two years six months there occurs Adam's first use of a conjoined reply in response to an adult sentence.

Adult: Well, here, come here and cut your paper.

Adam: I come here a cut it.

Now, how can the model be modified to produce such an utterance? In Grammar I, "I" could be combined with "come" to produce "I come." "Come" could be followed by a slot filled by "here" producing "come here," and "cut" could be followed by "it" meaning "paper" to produce "cut paper" or "cut it." The first model could produce the response "I come here cut it" by forming the relations "I come," "come here," and "cut it." "I come come here cut it" was formed by concatenating these three relations and "I come here cut it" was formed by a deletion rule. There was no way in which "I come here a cut it" could be produced.

We tried to use Model I to conjoin the single words "come" and "cut" by adding "and" to the lexicon and adding the concept of coordination to the concept set. Grammar I rules yielded

I come come here come and cut cut it

which the deletion rules transformed into

I here come and cut it.

This was clearly wrong. No child would make such an error. It seemed that either we must devise different deletion rules for different situations by distinguishing between at least subject and object so that conjunction reduction could be employed, or alternatively that we must use the phrasal approach. Either way the data forced us to form a new template rule which conjoined relation pairs, not merely relation words or slot fillers. Thus this rule is the first hierarchical rule in a previously flat grammar. Moreover since this rule provided us with an alternative way to fill slots, the Grammar II is now more conveniently represented by a set of phrase structure rules than by the flat templates. We have given only one illustrative sentence here due to constraints of space but the arguments presented here are supported by many sentences in the data.

Since it is not a straightforward task to transform the template grammar to distinguish between subject and object we found that it was easier to form Model II by adding the ability to form conjoined phrases and use them as slot fillers. It appeared that this simple strategy sufficed for the Adam data. We cannot be certain of the validity of this strategy unless we painstakingly develop Model II by permitting it to evolve week by week from language samples of a single child as we did with Model I. It is an interesting fact about the model that, however, that the conjoined phrase strategy is much simpler to implement and is apparently adequate.

In summary, all the new coordinated responses could be accommodated by adding a new syntactic word class to the model and permitting hierarchical structure represented by the addition of a context-free phrase structure rule which took the Model I template

relation-word <u>slot-filler</u>

and called it simply relation, e.g.,

relation --> relation-word <u>slot-filler</u> and then added the rule

slot-filler --> relation "and" relation

The conjunction "and" was inserted as the first entry in the closed class in the Model II lexicon. Note that this additional phrase structure rule causes Grammar II to be recursive as well as hierarchical.

One interesting aspect of this case study of the addition of conjunctions to Model I which was flat, and which resulted in a hierarchical and recursive grammar for Model II, is that the machinery forced by the data

was contrary to that which we had anticipated when we first started to explore the addition of conjunctions. Since the model already had concatenation and deletion rules, before experimenting with the model we assumed that the conjunction reduction approach would be easier to implement. Experimenting with the model and comparing it to the data forced us to the realization that the current concatenation and deletion rules would not work for conjoined sentences and neither could they be easily modified.

In conclusion we would emphasize that we take the position that there is no need to assume that the hierarchical grammar evolves all at once. Neither is there any need to assume that hierarchy is triggered by conjunctions for all children. Mechanisms may appear in a different order in different children's grammars. The interesting observations offered by this case study are simply that experimenting with a computational model will often surprise the designer, will suggest new questions to explore, and will provide new insights which may complement those of traditional linguistic analysis. What can we conclude about the conjunction reduction hypothesis vs. the conjoined phrase structure hypothesis from this case study? Certainly the model cannot be regarded as offering conclusive evidence for one choice over the other. It is, however, interesting to note that in experimenting with this model it was far easier to produce sentences such as Adam's first conjoined sentences by employing the conjoined phrase analysis rather than the conjunction reduction approach, and, as anticipated, the addition of "and" to the grammar forced the development of a hierarchical grammar.

References

Arbib, M. A., Conklin, E. J., and Hill, J. C., to appear, From Schema Theory to Language, Oxford University Press.

Brown, R. A., 1973, A First Language: The Early Stages, Cambridge Mass., University Press.

Hill, J. C., 1984, "Combining Two Term Relations" Evidence in Support of Flat Structure," Journal of Child Language, 11.673-678.

_____, 1983. "A Computational Model of Language Acquisition in the Two-Year-Old," Cognition and Brain Theory, 6(3).287-317.

_____, 1982, "A Computational Model of Language Acquisition in the Two-year-old," Ph.D. Dissertation, University of Massachusetts at Amherst, reproduced by the Indiana University Linguistics Club, Bloomington Indiana, February 1983.

Hill, J. C., and Arbib, M. A., 1984, "Schemas, Computation, and Language Acquisition," Human development, 27.282-296.

Lust, B., 1977, "Conjunction Reduction in Child Language," Journal of Child Language, 4. 257-87.

Lust, B. and Mervis, C., 1980, "Development of Coordination in the Natural Speech of Young Children," Journal of Child Language, 7.279-304.

Piattelli-Palmarini, M. (ed.), 1980, Language and Learning: The Debate Between Jean Piaget and Noam Chomsky, Cambridge, Mass., Harvard University Press.

Stockwell, P., Schacter, P., and Partee, B., 1973, The Major Syntactic Structure of English, Holt, Rinehart and Winston, N. Y.

Tager-Flusberg, H., deVilliers, J., Hakuta, K., 1982, "The Development of Sentence Coordination, " in Language Development: Problems, Theories and Controversies, Volume I: Syntax and Semantics, edited by S. Kuczaj II, Lawrence Erlbaum Associates, Hillsdale, N. J.