Sheldon Nicholl
Dept. of Computer Science
University of Illinois
1304 W. Springfield Ave.
Urbana, IL 61801

David C. Wilkins
Dept. of Computer Science
University of Illinois
1304 W. Springfield Ave.
Urbana, IL 61801

Abstract

This paper presents the result that a computer program can mimic the acquisition by children of a selected set of grammatical morphemes. Roger Brown [Brown, 1973] studied the acquisition of 14 morphemes, and showed how a set of partial order relations describes this aspect of child language learning. We show that these relations can be given a computational basis. They follow directly from a class of Boolean learning algorithms which have three simple constraints in the manner in which they consider hypotheses. I will call these three constraints the CAM constraints. CAM constraint 1 is to increase the length of the conjuncts one term at a time. The second CAM constraint is to consider all hypotheses of the same length simultaneously. Finally, CAM constraint 3 is to collect all single-term hypotheses involving noun features into a single conjunction prior to Boolean learning.

1 INTRODUCTION

The problem of how language is acquired is still one of the core problems in understanding cognition. Contributions to an understanding of language acquisition have come not only from Psychology and Linguistics [Brown, 1973], [Pinker, 1984] but also from Computer Science [Gold, 1967] and artificial intelligence [Berwick, 1985]. Therefore we would expect further progress in the study of language acquisition to draw more and more from multiple fields in the cognitive sciences, rather than to draw from purely one area.

Brown's results [Brown, 1973] provide a fine example of this interdisciplinary transfer. In studying the language development of three children, Brown focussed his attention on the acquisition of 14 particular grammatical morphemes. Brown was able to write down the acquisition order of a significant subset of these grammatical morphemes in terms of partial orders. This still stands as a major result in the study of language acquisition.

The goal of this paper is to exhibit a computer model of some of Brown's acquisition order results (see Table 3). This computer model is based on a set of constraints for whose plausibility we hope to argue. I will call these constraints the CAM constraints. The CAM constraints are part of a larger model of language acquisition called CAM (Categories, Agreement, and Morphology) which has been implemented in Common Lisp.

2 BROWN'S RESULTS

Brown used the idea of a complexity ordering to lend an underlying paradigm to his order-of-acquisition data. Different grammatical constructions are ranked on a scale of increasing complexity; the more complex the construction, the longer it takes to learn. Brown considered two complexity orderings, one based on syntax, the other based on semantics. The syntax-based ordering is stated in terms of a theory which is seriously out of date (Chomsky's Standard Theory [Chomsky, 1965]) and also depends on transformations, which CAM does not represent; this ordering will therefore not be further considered here.

What is of interest here is Brown's semantic complexity ordering. It is appropriate to begin with a description of the actual semantics that form the basis of the ordering.

Table 1 shows the meanings of the plural and past morphemes.[1] Note that although Table 1 refers to both regular and irregular forms, this distinction is not of consequence here, since both forms encode the same meaning. This is true of the past forms as well as the third person forms. Now the third person forms refer

[1] As Brown points out, the subjunctive use of the past which is used to refer to unactualized events, need not be considered here, since the children learn this meaning of the past long after the morphemes of interest here have already been learned.