J. Child Lang. **31** (2004), 959–961. Printed in the United Kingdom

DISCUSSION

DOI: 10.1017/S0305000904006452

Gold's theorems and the logical problem of language acquisition

BARBARA C. SCHOLZ
Philosophy Department, San Jose State University, and Linguistics Research Center, University of California Santa Cruz
scholz@ling.ucsc.edu

Evaluating Brian MacWhinney's multiple 'solutions' to the logical problem of language acquisition requires delineating exactly what the alleged problem is. He takes it to stem from the theorems of Gold (1967), the most celebrated of which is G1.

G1 No class of languages that is super-finite (i.e. contains all finite languages and at least one infinite language) is identifiable in the limit from text.

G1 entails the falsity of MacWhinney's claim that 'we know that finite-state grammars can be acquired from positive evidence.' (Every finite language is finite-state, and some finite-state languages are infinite, hence the class is not identifiable from text.) I assume he meant FINITE languages.

G1 poses a challenge to psychological theories of language acquisition only if all of the assumptions of Gold's paradigm hold for actual L1 learners: that input to actual learners of a language L is a text for L ('a sequence of strings x_1, x_2, … from L such that every string of L occurs at least once'; Gold, 1967: 450); that languages are learned by hypothesis testing; that success is identification in the limit; that what is learned is a generative grammar; and crucially, that the class of natural languages is superfinite. Under those conditions, G1 entails that learning natural languages is impossible. Yet children do learn them – contradiction.

This paradox can be avoided by denying that Gold's assumptions apply to L1 learners. We could, for example deny any of these:

(i) that languages are learned by means of hypothesis formation and testing, or
(ii) that the input to the child learner is merely text, or
(iii) that what is learned is a generative grammar that exactly generates the target language.

The principles-and-parameters approach to explaining language acquisition rejects (i). Principles and parameters are not hypotheses that are tested. The statistical learning theory of Horning (1969), Charniak (1993) and Haussler (1996) – mentioned by MacWhinney (2004: 11) but not pursued – rejects (ii). Probabilistic studies are entirely outside the paradigm Gold defines. And Pullum & Scholz (2003: 135–9) reject (iii). They reject the equation between learning a language and identifying a generative grammar that exactly generates it. Rejecting any of Gold's assumptions evades the paradox by denying the relevance of G1 to real learners.

Curiously, although MacWhinney mentions this road, the one he takes is different. He proposes that learning is CONSERVATIVE. It is important to see that the paradox is untouched by this suggestion. G1 shows that given Gold's assumptions there is NO effective strategy that guarantees successful identification in the limit from text for a super-finite class. That includes conservative strategies.

To see why, contrast two strategies for learning adopted by the imaginary learners Bold Bonnie and Cautious Connie introduced in Pullum & Scholz (2003: 129–35). Bold Bonnie will under some conditions hypothesize grammars for infinite languages. If she ever hypothesizes a grammar for an infinite proper superset of a finite target language, she can never recover, since no text can refute her over-liberal hypothesis. This is the side of the proof that MacWhinney talks about in terms of 'over-generalization', although it has nothing to do with generalizations based on word forms, but only those about the cardinality of the target language. By contrast, Cautious Connie never hypothesizes a grammar for an infinite language if some grammar for a finite language is consistent with the sequence of strings presented so far. In a super-finite class there will always be one, so if the target language is infinite, Cautious Connie will forever hypothesize grammars for successively larger finite languages, and thus never succeed. This conservative strategy guarantees failure in Gold's terms. So neither Bold Bonnie's 'over-generalizing' strategy nor Cautious Connie's conservative strategy can succeed. Gold showed that EVERY strategy has either Bonnie's problem or Connie's.

MacWhinney claims that inferred negative evidence provides another solution to the logical problem. Here he is relying on an entirely different series of Gold's theorems, not relating to text. One example is G2.

G2 The class of primitive recursive languages is identifiable in the limit using information presentation by an informant.

The proof depends on the same assumptions as for G1, except that the learner is assumed to be exposed, in the limit, to the entire set of facts about both what is and what is NOT in the target language – not just some of the facts but *all* of them. We hardly need the empirical findings of Brown & Hanlon (1970) to show that the primary linguistic data (PLD) never provides a complete presentation of the complement of the language.

MacWhinney's proposal is that learners might INFER from the absence of some word form or construction X in the PLD that X is ungrammatical in the target language. But this does not show that G2 applies to child learners. First, it doesn't show that learners infer a complete enumeration of the complement of the target language. Second, inferring the absence of strings in the target language from any absence in the PLD would undermine learning rather than assist it. (To see this, suppose a learner has never heard *wombat* and *soliloquy* in the same clause, and concludes that their co-presence is grammatically forbidden. The conclusion is false. In any corpus that does not exhaust the target language there will be indefinitely many such grammatically irrelevant absences of evidence. So a learner who infers absence from the target language on the basis of absence in the PLD will hypothesize an unending series of incorrect grammars and will never succeed in Gold's terms.) And third, it is inconsistent to defuse the paradox of G1 by denying the applicability of Gold's assumptions to real learners, and then turn around and claim they DO apply to real learners in order to help oneself to the positive result G2.

Conservative learning strategies, and inference from absence of evidence to evidence of absence, may solve some problems and puzzles about item-based language acquisition. But they do not solve the paradox based on G1. Nor does the possibility of inferring some negative evidence from its absence in the PLD show that G2 applies to real language learners.

REFERENCES

Brown, R. & Hanlon, C. (1970). Derivational complexity and order of acquisition in child speech. In J. R. Hayes (ed.), *Cognition and the development of language*. New York: Wiley.

Charniak, E. (1993). *Statistical language learning*. Cambridge, MA: MIT Press.

Gold, E. M. (1967). Language identification in the limit. *Information and Control* **10**, 447–74.

Haussler, D. (1996). Probably approximately correct learning and decision-theoretic generalizations. In P. Smolensky, M. C. Mozer & D. E. Rumelhart (eds), *Mathematical perspectives on neural networks*. Mahwah, NJ: Erlbaum.

Horning, J. (1969). *A study of grammatical inference*. Stanford University, Computer Science Department.

Pullum, G. K. & Scholz, B. C. (2003). Linguistic models. In M. Banich & M. Mack (eds), *Mind, brain, and language: multidisciplinary perspectives*. Mahwah, NJ: Erlbaum.