J. Child Lang. **31** (2004), 941–943. Printed in the United Kingdom
© 2004 Cambridge University Press

DISCUSSION

DOI: 10.1017/S0305000904006415

Comment on 'A multiple process solution ...' (B. Macwhinney)

PARTHA NIYOGI

The University of Chicago

The central question posed by the so called 'logical problem of language acquisition' is how it comes to be that children are able to GENERALIZE from a finite set of linguistic data to acquire (learn, develop, grow) a computational system (grammar) that applies to novel examples not encountered before. The difficulty of this generalization problem was first posed cogently by Gold and while Macwhinney discusses the Gold framework and the linguistic literature on this matter, it is worth noting that the Gold framework is not the only one. There are at least two important new sources of insight from computational learning theory in the decades following Gold that need to be kept in mind. First, there is the development of empirical process theory that forms the basis of any analysis of statistical learning (see summary in Vapnik, 1998). Applying this approach to language (see Niyogi, 1998 for a treatment), one concludes that the family of learnable grammars must have a finite Vapnik Chervonenkis (VC) dimension. The VC dimension is a combinatorial measure of the complexity of a class of functions. Grammars may be viewed as functions mapping sentences to their grammaticality value. In this more sophisticated sense of the VC dimension, the class of grammars must be constrained. Second, there is the development of the theory of computational complexity suggesting that while a learning algorithm might exist, it may not be efficient, i.e. run in polynomial time. These two developments come together in the influential Probably Approximately Correct (PAC) model (Valiant, 1984).

To illustrate the power of these insights consider the case of learning finite state grammars (regular languages). It turns out that these have infinite VC dimension and therefore are unlearnable from positive and negative examples in the statistical PAC sense. In contrast, regular languages are learnable in the Gold tradition (from positive and negative examples). For that matter, even the class of all finite languages are learnable in the Gold setting (in fact, from positive data alone) but unlearnable from positive and negative examples in the PAC setting. In other words, the arguments for constrained grammar families come from many different theoretical traditions in computer science and statistics.

https://doi.org/10.1017/S0305000904006415 Published online by Cambridge University Press

One will therefore need to consider Macwhinney's solutions in the light of this larger body of literature. I will comment briefly on three of his solutions to clarify some points that need to be properly understood.

1. Limiting the class of grammars. It is important to recognize upfront that this is exactly the solution offered by the theory of universal grammar. Any posited constraint on the class of grammars may be judged by two criteria – learnability and interesting linguistic or cognitive insights that may be derived from it. The onus then falls on the researcher to demonstrate interesting insights obtained by the formalism and its implied constraints.

2. Conservatism. It is important to note that conservatism by itself does not avoid Gold's theorem. Recall that Gold's theorem is about the informational inadequacy of ANY learning algorithm – in particular, a conservative learner following some sort of simplicity principle is also subject to it. Nevertheless, conservative learning is a powerful idea and occurs in many different forms across a variety of approaches. For example, Chomsky's early notion of an evaluation metric on grammars is also an instantiation of this idea.

3. Item-based learning. This sounds awfully like taking the view that language is just a set of lists. While one may argue that arbitrary recursion is impossible because of memory constraints and therefore that all languages are ultimately finite, one usually notices a combinatorial explosion in the number of possible sentence types. The list based or item based approach seems suspect for such combinatorial reasons.

Ultimately, any learning algorithm may be viewed as a map $\mathcal{A}$: Data$\rightarrow\mathcal{G}$ where *Data* is the space of possible linguistic experience and $\mathcal{G}$ is the space of possible grammars. Thus, given any particular data set of linguistic examples ($d \in Data$), the algorithm acts on it to produce a grammar, i.e. $\mathcal{A}(d)$ is a grammar belonging to $\mathcal{G}$. By learnability arguments, in a variety of different frameworks, one is forced to conclude that $\mathcal{G}$ is clearly not the set of all phrase structure grammars. It is a more restricted subset and in this sense, something like universal grammar must be true. Principles and parameters is an approach to providing a low dimensional characterization of $\mathcal{G}$ (metaphorically much like principal components analysis to those familiar with data analysis).

However, that does not at all mean that this restricted family $\mathcal{G}$ *does* admit a simple mathematical characterization. On the other hand, it may turn out to be that the algorithm $\mathcal{A}$ may admit an easy characterization while its image (the space of grammars) may not. If this were the case, it would be fruitful to provide a constructive description of the learning algorithm without ever being able to describe the class of grammars that are being learned. In this sense, MacWhinney's philosophical position may have some merit though the details are far from worked out at the current time.

https://doi.org/10.1017/S0305000904006415 Published online by Cambridge University Press

REFERENCES

Niyogi, P. (1998). *The informational complexity of learning: perspectives of neural networks and generative grammar*. Kluwer Academic Publishers.
Valiant, L. (1984). *A theory of the learnable*. Communications of the ACM.
Vapnik, V. (1998). *Statistical learning theory*. John Wiley.

https://doi.org/10.1017/S0305000904006415 Published online by Cambridge University Press