Brit. J. Phil. Sci. **32** (1981), 157–176 *Printed in Great Britain*

Discussions

CAN POPPERIANS LEARN TO TALK?

In several recent publications (Sampson [1978], [1980*a*]) Geoffrey Sampson has argued that an essentially Popperian language acquisition device could learn language much as a human child does. The device Sampson envisions would freely (or perhaps randomly) generate hypotheses about the grammar the child seeks to learn, and test these hypotheses against the data available to the child. If the data are incompatible with an hypothesis, the hypothesis is rejected and another one tried. If any hypothesis does not conflict with the data, it is retained for further testing. Sampson's language acquisition device exhibits a Popperian proclivity for strong, simple hypotheses, but when it sets out on its acquisition task it has no *a priori* information about the nature of the grammar it seeks to acquire. If Sampson is right, then there is no need to postulate rationalist learning mechanisms of the sort advocated by Noam Chomsky (in Chomsky [1975] and many other publications) to account for the child's ability to learn language. However, it is my view that Sampson's argument is fatally flawed. He has not shown that a Popperian learning mechanism could duplicate the child's feat. Indeed, I think it can be shown that Sampson's Popperian learner could not possibly match the language learning achievement of the normal child, save by miraculous accident. In an earlier paper (Stich [1979]) I set out the argument that a Popperian device could not learn language, and I explained where Sampsons defence of the Popperian learning mechanism had gone astray. My conclusion was *not* that Chomsky's rationalist account of language learning must be correct, for I see serious problems with that view as well. Rather, I urged that the correct theory of language acquisition ought to be sought in the very substantial domain that lies between Chomskian rationalism and Popperian empiricism.

Sampson was not convinced. In a reply to my paper (Sampson [1980*b*]) he argues that I am impaled on the horns of a dilemma: if one of the premises of my argument is true, then I could not possibly know the other to be true. What I shall argue in the present note is that Sampson's attempt to rescue his Popperian account of language acquisition does not succeed. To make my case, I will begin with a brief sketch of the argument which shows that Sampson's Popperian learning mechanism could not duplicate the child's achievement. Next I will attend to Sampson's alleged dilemma, giving it the most charitable reading I can come up with, and showing why even on this charitable reading the dilemma fails.

1 A SKETCH OF THE ARGUMENT AGAINST SAMPSON'S POPPERIAN LEARNING MECHANISM

The argument I use to show the inadequacy of Sampson's Popperian learning theory is one that I have labelled *The Rational Scientist Argument*. Only the label is mine, however; the basic idea of the argument is due to Chomsky. Stripped to its essentials, the Rational Scientist Argument requires an assumption and a pair of premises. The assumption is that when a person has learned a language he has acquired a tacit knowledge of the rules of its grammar, and that he invokes this tacit knowledge (or tacit theory) when he produces, interprets and makes judgements about sentences. Sampson and I agree that there are 'large conceptual problems' with this assumption.[1] But we are both prepared to grant that 'these problems are not important for the matters discussed below'.

The *first premise* of the Rational Scientist Argument claims that if we were to present a rational scientist with the data available to a typical human child and set the scientist the chore of discovering the grammar tacitly known by the child's senior co-linguists, the rational scientist would fail. Given only the data available to the child, the scientist could not discover the grammar of the child's linguistic community, save by accident. The premise itself requires an argument, of course. The basic point in the argument for the first premise is that there will be indefinitely many grammars compatible with all the data available to the child. Many of these grammars will differ substantially from the grammar of the child's elders. In trying to decide among the many grammars compatible with his data, the scientist may invoke various methodological criteria such as simplicity or strength. However, there is little reason to believe that such methodological criteria would be strong enough to select a single grammar. And if they were, there is no reason to think that the single grammar selected would much resemble the grammar of the child's linguistic community. Clearly, this bare bones version of the argument for the first premise is open to attack on several fronts. But its problems need not be ours. For Sampson is ready to concede, if only for argument's sake, that the first premise is true.

> To my mind the wealth of linguistic data available to a child, the tendency to choose strong and simple theories . . ., and the constraints implied by Simon's argument, will between them narrow down the set of possible grammars to a class whose members will differ from one another rather little. . . . But I do not need to establish this in order to counter Stich; *I can allow, as he suggests, that the factors just listed may conjointly be compatible with grammars which differ widely with respect not merely to their internal structure but to the predictions they make about observable behaviour.* (Emphasis added.)

[1] The quote is from Sampson [1980*b*]. All subsequent quotes will be from this paper, unless otherwise indicated. For my qualms about the tacit knowledge assumption, see Stich [1971].

The *first premise* has a pair of interesting corollaries. First, since the imagined rational scientist can invoke all of the strategies of discovery and inference dreamt of in empiricist theories of mind, it follows that no learning mechanism restricted to empiricist strategies of inference and discovery could reliably discover the grammar of the child's linguistic community, given only the data available to the child. A second corollary is a special case of the first. Since the Popperian strategy of free conjecture and refutation is clearly among those dreamt of by empiricists, it follows that Sampson's Popperian learning mechanism could not discover the grammar of the child's elders, given only the data available to the child.

The *second premise* of the rational scientist argument is that the child *does* tacitly come to know much the same grammar as is tacitly known by his senior co-linguists. Since the dispute between Sampson and myself turns largely on this premise I had best elaborate on just how I intend the premise to be understood. There are many ways in which the grammars tacitly known by a pair of people might resemble one another. The strongest relation that may obtain between a pair of tacitly known grammars is *strict identity*—the grammars contain all and only the same rules in the same order. A weaker relation between a pair of grammars is what might be called *strong generative equivalence*. A pair of grammars are strong generatively equivalent if they generate the same class of sentences and assign them the same structural descriptions. A still weaker relation among grammars is what I shall call *weak generative equivalence*. Here what is required is merely that the grammars generate the same class of sentences (whether or not they assign them the same structural descriptions). Finally, a pair of grammars may be what I shall call *roughly equivalent*. For this notion I have no precise definition. A pair of grammars are roughly equivalent if they are reasonably close to being weak generatively equivalent; grammars are roughly equivalent when the classes of sentences they generate *substantially* overlap. So characterized, rough equivalence admits of degrees. A pair of weak generatively *equivalent* grammars would be at the very top of the rough equivalence scale, while the grammar of English I tacitly know and a grammar tacitly known by a monolingual Korean would be well down toward the bottom. Note that strict identity entails strong generative equivalence which entails weak generative equivalence which entails rough equivalence, though in no cases does the opposite entailment hold. Now when the second premise of the Rational Scientist Argument claims that the child comes to know tacitly 'much the same' grammar as is tacitly known by his elders, I would have it understood as claiming that the child comes to know tacitly a grammar which is *roughly equivalent* to those known by his senior co-linguists.[1]

[1] We might also define a notion of *rough-strong-equivalence* which is related to strong generative equivalence as rough equivalence is related to weak generative equivalence. That is, a pair of grammars are rough-strong-equivalent if and only if they generate roughly the same class

What the premise is claiming, then, is that the class of sentences generated by the child's grammar is very nearly identical with the classes of sentences generated by the grammars of his elders.

With this explanation of rough equivalence behind us, we can now retrace our steps and state the first premise of the Rational Scientist Argument a bit more precisely. As stated three paragraphs back, the first premise claimed that the rational scientist could not discover *the* grammar known by the child's senior co-linguists, given only the data available to the child. Cast in this way the first premise must presuppose that there is a *single* grammar known by all the child's seniors, a presupposition which is more than a little implausible. What is much more plausible is that the child's elders (or at least a very substantial majority of them) tacitly know grammars which are roughly equivalent to each other. What the first premise must claim, then, is that the rational scientist whose only data are the data available to the child could not come up with a grammar roughly equivalent to those of the child's elders, save by chance. This sharpening of the first premise should not be problematic for Sampson, since he was prepared to grant that the child's data along with all appropriate methodological constraints 'may be conjointly compatible with grammars which differ *widely* with respect not merely to their internal structures but to the predictions they make about observable behaviour'. (Emphasis added.)

Given the premises of the Rational Scientist Argument, the conclusion follows straightforwardly. Since the child does what the rational scientist cannot do, the child must have available to him information or inferential strategies not available to the rational scientist. Were the child to be equipped with a Popperian mind, it could not learn to speak.

2 SAMPSON'S DILEMMA

Let us turn to Sampson's dilemma. The core of his argument is given in the passage which follows the long quote reproduced above.

> All I need do, essentially, in order to defeat Stich is to ask: *How does he know* that children invariably 'get the right answer' when learning their elders' language?
>
> The only way one *could* know this would be to act as a rational scientist by formulating theories about the grammars of a number of individual members of a language-community, in order then to show that these grammars were very similar to one another, or to the grammars of members of the previous generation (which would also have to be discovered by acting as a rational scientist). Now Stich may play the game whichever way he wishes. If he says that the rational scientist will have no

---

of sentences and assign to them roughly the same structural descriptions. In Stich [1979] I ran the Rational Scientist Argument using the notion of rough-strong-equivalence rather than rough equivalence, though I did not pause to give the relation a name. *Cf.* §2. For the purposes at hand, the difference makes no difference, save that the argument is briefer and neater if it uses rough equivalence in Premise 2.

means of choosing between widely different grammars on the basis of the kind of data-set available to a child, then I reply that the rational scientist will in that case certainly not be in a position to state that the grammars of two individuals are near-identical, since no data are available to the rational scientist that would not be available equally to a child. If, on the other hand, Stich says that the rational scientist *can* ascertain the nature of individuals' grammars precisely enough to confirm that they are closely similar, then I reply that in that case the child can ascertain the precise nature of the elders' grammar by the same technique.

With a bit of rational reconstruction, we can portray Sampson's argument as an explicit dilemma:

Either (*A*) a rational scientist using the data available to a child *can* determine the grammar(s) of the child's elders, or (*B*) a rational scientist using the data available to the child *cannot* determine the grammar(s) of the child's elders. But (*A*) is a direct denial of Premise 1, so if (*A*) is true, then the Rational Scientist Argument simply has a false premise. If we opt for (*B*), however, then we could not possibly know that Premise 2 is true.

But now *why* does (*B*) entail that we could not possibly know that Premise 2 is true? Sampson's argument for this crucial entailment is, unfortunately, very compressed. He says that if (*B*) then 'the rational scientist will in that case certainly not be in a position to state that the grammars of two individuals are near-identical, since no data are available to the rational scientist that would not be available equally to the child'. Construed in what I take to be the most charitable way, Sampson is arguing as follows:

(*i*) If a rational scientist were to try to discover an individual's grammar *with no restrictions placed on the sort of data he could use*, he would as a matter of fact find himself trying to discover the grammar on the basis of the same sort of data that might be available to a child, since 'no data are available to the [unrestricted] rational scientist that would not be available equally to the child'.

(*ii*) from (*i*) and (*B*) it follows that the unrestricted rational scientist could not determine the individual's grammar.

(*iii*) If the unrestricted rational scientist cannot determine the grammars of individual members of a language community, then he cannot know that the grammars of these individuals are near identical

(*iv*) From (*ii*) and (*iii*) it follows that an unrestricted rational scientist could not know that the grammars of individual members of a language community are near identical.

(*v*) If the unrestricted rational scientist could not know this, we could not know it either. So we cannot know that Premise 2 is true.[1]

[1] In uncharitable moments the thought crosses my mind that Sampson's argument is not concerned with what an 'unrestricted' rational scientist could do. On this uncharitable interpretation, (*i*) and (*ii*) would be replaced by

If in (*i*)–(*v*) I have succeeded in capturing the argument Sampson has in mind, then there is a pair of objections *each* of which, I think, suffices to show that Sampson's argument is unsound, and thus that his proposed dilemma inflicts no wounds on the Rational Scientist Argument. The first objection focuses on step (*i*) of Sampson's argument, the second on step (*iii*). Let me consider them in turn.

Is it the case, as Sampson seems to claim, that a rational scientist seeking to discover an individual's grammar with no restrictions on the data he may use would in fact have no useful data available that is not available equally to the child? The answer, I think, is clearly *no*. Linguists can and do solicit from native speakers intuitions about grammaticality, about various syntactic properties and relations and about various semantic properties and relations. *Without using such intuitions as data, modern generative grammar could hardly have gotten started.*[1] But there is no reason to believe that the child has access to any significant cache of data about the intuitive judgements his elders render concerning the syntactic and semantic features of their language. In addition, linguists and psycholinguists can and do collect data on a variety of other phenomena in an effort to decide among competing hypotheses about the grammar of a community. Latency times, relative frequencies of slips of the tongue, the differences between sounds actually on a tape and what a subject reports hearing, interference effects, and more may all prove useful to the rational scientist attempting to uncover the details of a grammar. *But no such data are available to the child.* So it would appear Sampson is just *wrong* in claiming that no data are available to the scientists which are not available to the child.

Even if Sampson were right on this point, however, his argument would still be fatally flawed. For in step (*iii*) the argument claims that a scientist can determine a pair of tacitly known grammars are near identical only by determining what the grammars are.[2] But this is simply false.

(*i–ii*) If a rational scientist were to try to discover an individual's grammar, and if his data were so restricted that 'no data are available to [him] that would not be available equally to the child', then from (*B*) it follows that this restricted rational scientist could not determine the individual's grammar.

In the remaining steps, 'unrestricted' would be systematically replaced by 'restricted'. However, on this uncharitable construal, step (*v*) becomes wildly implausible and is left without any defence.

[1] Sampson has objected to this use of intuitions in modern linguistics. *Cf.* Sampson [1979], pp. 186 *ff*. I find his objections utterly unpersuasive, though I shall not pursue the matter here. For my account of the role of intuitions as data for generative grammar, *cf.* Stich [1971] & [1972].

[2] Sampson's claim is actually a tiny bit weaker. According to him, 'The only way we could know [that children invariably get the right answer when learning their elders' grammars] would be to act as a rational scientist by formulating theories about the grammars of a number of individual members of a language community, in order then to show that these grammars were very similar to one another. . . .' However, the difference between this claim and (*iii*) does not affect the argument to follow.

To know that a pair of tacitly known grammars are roughly equivalent in the sense required by Premise 2 of the Rational Scientist Argument, *we need know nothing whatever about what the rules of the grammars actually are*. Recall that to say a pair of grammars are roughly equivalent is simply to claim that the classes of sentences they generate substantially overlap. And we can have overwhelming inductive evidence for this claim without having any theory whatever about the rules of the grammar. If a pair of speakers generally find each other's utterances comprehensible, this is some evidence that the grammars they tacitly know are roughly equivalent. Moreover, if the sentences uttered by one speaker rarely or never strike the second speaker as deviant, and if the sentences uttered by the second speaker rarely if ever strike the first as deviant, this is still better evidence that the grammars tacitly known by the two speakers are roughly equivalent, whatever their rules may be. The strength of such inductive evidence increases along with the number of sentences uttered by each speaker in the presence of the other. Thus we have overwhelming inductive evidence that our friends and children tacitly know grammars roughly equivalent to our own, even if we have formulated no theories at all about our grammar or theirs.[1]

We may drive home the point by considering a pair of analogies. It is no easy matter to infer the program of a hand held electronic calculator by observing its inputs and outputs, the time it takes to do various calculations, *etc*. It is much easier to discover that the programs of a pair of calculators are roughly input–output equivalent. If we do a few thousand varied calculations on each, and discover that given identical input they yield identical output, this is strong evidence that their programs are roughly input–output equivalent. And it is evidence we can acquire even though we haven't the vaguest idea what the program of either calculator is.

A hand held calculator, of course, is not a terribly good analogy for a speaker. So consider a second, closer, analogy. Suppose we have a computer programmed to make acceptability judgements about sentences. Given a sentence, the computer will judge it to be either acceptable, unacceptable or partially acceptable (perhaps to some degree). Notoriously, it would be no simple feat to program a computer which would come close to duplicating the acceptability judgements of a native speaker of

[1] We also have an enormous amount of indirect evidence bearing on the claim that a given child's grammar is roughly equivalent to that of his senior co-linguists. For we know that it is generally the case with other children that their grammars are roughly equivalent to those of their seniors. And there is evident selective advantage to a psychological mechanism which guarantees that the child will acquire a grammar roughly equivalent to those in his linguistic environment. Thus, in the absence of any evidence that a given child or his situation is abnormal, it would be reasonable to assume that the child's grammar is roughly equivalent to those of his elders, even before accumulating any substantial amount of evidence about how well the child and his elders understand each other's utterances.

English. Suppose, however, that someone has succeeded in programming a computer which does a passable job at mimicking, say, my acceptability judgements. Now suppose we are presented with a second computer which also has been programmed to make acceptability judgements. And suppose the question arises whether the second computer is roughly equivalent to the first, in the sense that it marks roughly the same classes of sentences as acceptable, unacceptable or partially acceptable. Sampson's premise (*iii*) seems to suggest that we could not determine whether the two computers are roughly equivalent without determining the program of each. But this is simply false. We can have overwhelming evidence that the two computers are roughly equivalent, even though we haven't the vaguest idea what program is being run by either. Analogously, we can (and do!) have overwhelming evidence that our children and neighbours have grammars roughly equivalent to our own, though we (or at least I) have no serious idea what the neighbour's grammar or our own may be.

Where does all this leave us? I have been arguing that, for two quite independent reasons, Sampson's argument as reconstructed in (*i*)–(*v*) is unsound. Thus it does not follow that if we opt for horn (*B*) of his dilemma we could not know Premise 2 of the Rational Scientist Argument. I conclude that the horns of Sampson's dilemma leave the Rational Scientist Argument unscratched.

STEPHEN P. STICH
The University of Maryland, College Park

REFERENCES

CHOMSKY, N. [1975]: *Reflections on Language*. New York: Pantheon Books.
SAMPSON, G. [1978]: 'Linguistic Universals as Evidence for Empiricism', *Journal of Linguistics*, **14**, pp. 183–206.
SAMPSON, G. [1979]: *Liberty and Language*. Oxford University Press.
SAMPSON, G. [1980*a*]: *Making Sense*. London: Cambridge University Press.
SAMPSON, G. [1980*b*]: 'Popperian Language Acquisition Undefeated', *British Journal for the Philosophy of Science*, **31**, pp. 63–67.
STICH, S. [1971]: 'What Every Speaker Knows', *Philosophical Review*, **lxxx**, pp. 476–96.
STICH, S. [1972]: 'Grammar, Psychology and Indeterminacy', *Journal of Philosophy*, **lxix**, pp. 799–818.
STICH, S. [1979]: 'Between Chomskian Rationalism and Popperian Empiricism', *British Journal for the Philosophy of Science*, **30**, pp. 329–47.

CONSEQUENCES OF A SIMPLE EXTENSION OF THE DUTCH BOOK ARGUMENT

One of the main arguments used to establish subjective prior probabilities as an additive measure is the need to avoid a 'Dutch Book', *i.e.* the 'first