Communicative Events as Evidence in Linguistics[†]

ROBERT J. STAINTON
Carleton University

1. INTRODUCTION

1.1 The Game Plan

MY PRINCIPAL AIM IN THIS PAPER is to reply to the demand, often heard in Systemic Functional circles, that one supply 'naturally occurring language episodes' — rather than speaker's intuitions — to support linguistic hypotheses. I think this demand is, at best, generally unjustified; at worst I consider it a harmful distraction.[1]

By way of response to the demand, I will present arguments against the following claims:

(1) *The Strong Claim:* Texts are the only evidence in linguistics.

(2) *The Weaker Claim:* Texts are the best evidence in linguistics; (better, in any case, than native speaker's intuitions).

I think each of these is false, though my degree of certainty varies. I'm *positive* that the Strong Claim is wrong, and I have serious doubts about the Weaker Claim. About the Weakest Claim, given below, I remain agnostic.

(3) *The Weakest Claim:* Texts are very good evidence in linguistics.

[†] I would like to express my great debt to Michael Gregory, who first introduced me to issues in linguistic methodology and linguistic epistemology. I should note, however, that Michael is not entirely to blame for the views that I have come to hold; much of the fault on that score lies with Sylvain Bromberger and Noam Chomsky. Thanks also to Elissa Asp, Andrew Brook, Maria Casas, William Downes, Peter Fries, Nada Khirdaji and Gordon Tucker for valuable comments on an earlier draft.

1. Some will say the horse I'm about to beat has already "rung down the curtain and joined the choir invisible." Frankly, I agree. But the news of the horse's demise has been lamentably slow to spread.

1.2 Some Terminology

One can usefully think of linguists as attempting to answer a cluster of questions. One empirical question to which every linguist should want an answer is:

(4) Q: What is the grammar of the particular language L?

Obviously there are other questions which linguists also pursue. For instance: How do humans learn languages? How do humans use languages? What is the relationship between a group's language and its culture? And so on. But, I take it, these sorts of projects presuppose that linguists are able to find a grammar for L, the language used by some community — i.e. that linguists are able to arrive at some kind of answer to Q.

In particular, it seems to me that an adequate analysis of *a particular text* presupposes some kind of answer to Q, for the language of the text. Allow me to illustrate.

"A Birthday Party on Monday Night"[1]

<Ann> — Pardon . . . euh . . . C'est à quelle heure qu'on amène l'autre groupe ?
<Élyse> — Euh . . . mais . . . là . . . maintenant je crois.
<Russ> — Is today Monday? Yes it is unfortunately.
<Élyse> — Aujourd'hui c'est lundi.
<Russ> — Lundi.
<Élyse> — Quand est-ce que tu commences l'école ?
<Christian> — We have a birthday party Saturday and Monday and when . . . Then I go to family camp. A birthday party on Monday night.
<Élyse> — C'est ta fête ou la fête de quelqu'un d'autre?
<Christian> — It's a birthday party next week.
<Élyse> — Oui mais ta fête ?
<Christian> — Me.

Someone doing an analysis of the above text might like to point out things like:

1. This text is a transcription of a conversation which occurred at LE CAMP, a French immersion camp held at Glendon College, in August, 1986. The names are fictitious. For the entire text see Stainton and Hillier (1990: 97ff.).

a. Which sentences are in English (e.g. "It's a birthday party next week"), and which are in French (e.g. "Quand est-ce que tu commences l'école?")

b. Which expressions are perfectly grammatical ("C'est à quelle heure qu'on amène l'autre groupe?") and which are somehow irregular (e.g. "Lundi" is given an anglicized pronunciation by Russ, and Christian ends a sentence with "and when . . .")

c. Which expressions are formal ("unfortunately") and which informal ("Euh . . . mais . . . là . . . maintenant je crois"); and which utterances are too formal or informal, given the context (e.g. "unfortunately" as used by Russ sounds oddly formal)

d. Which utterances are congruent (e.g. "Quand est-ce que tu commences l'école?", an interrogative, is used by Elyse to ask a question) and which incongruent (e.g. "C'est ta fête", a declarative, is also used by Elyse to ask a question)

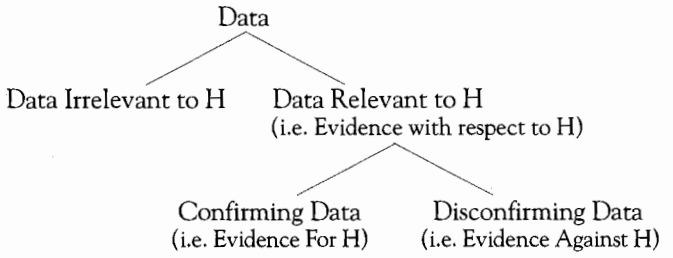
e. Which expressions are sentences (e.g. "Aujourd'hui c'est lundi"), and which are minor clauses (e.g. "Oui mais ta fête")

The list goes on. But notice: justifying these sorts of claims about the text may well require making recourse to two grammars: one for English and one for French. Hence, it seems fair to suppose, even linguists whose work is restricted to text analysis should want an answer to Q for the languages of the texts on which they work.

I will call tentative answers to scientific questions *hypotheses*, where a hypothesis can treat of some particular fact or event, or state a generalization. A *grammar* of L, being an answer to a scientific question (namely, Q) is a hypothesis; a hypothesis about L. (Or, if you prefer, a grammar of L is a related series of hypotheses about L.)

One tests a hypothesis against the *experiential data*, where this data seems to divide into two classes, for any hypothesis H: that data which is relevant to H (i.e. the *evidence* with respect to H) and that data which is irrelevant to H. That is:

(5)



a. Which sentences are in English (e.g. "It's a birthday party next week"), and which are in French (e.g. "Quand est-ce que tu commences l'école?")

b. Which expressions are perfectly grammatical ("C'est à quelle heure qu'on amène l'autre groupe?") and which are somehow irregular (e.g. "Lundi" is given an anglicized pronunciation by Russ, and Christian ends a sentence with "and when . . .")

c. Which expressions are formal ("unfortunately") and which informal ("Euh . . . mais . . . là . . . maintenant je crois"); and which utterances are too formal or informal, given the context (e.g. "unfortunately" as used by Russ sounds oddly formal)

d. Which utterances are congruent (e.g. "Quand est-ce que tu commences l'école?", an interrogative, is used by Elyse to ask a question) and which incongruent (e.g. "C'est ta fête", a declarative, is also used by Elyse to ask a question)

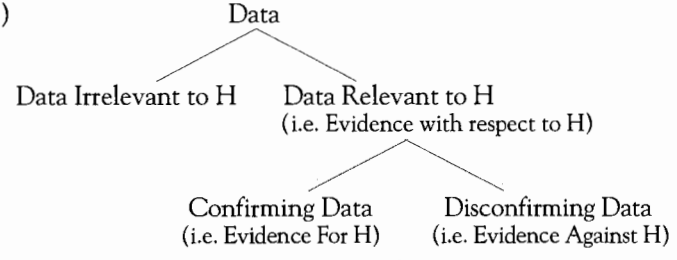
e. Which expressions are sentences (e.g. "Aujourd'hui c'est lundi"), and which are minor clauses (e.g. "Oui mais ta fête")

The list goes on. But notice: justifying these sorts of claims about the text may well require making recourse to two grammars: one for English and one for French. Hence, it seems fair to suppose, even linguists whose work is restricted to text analysis should want an answer to Q for the languages of the texts on which they work.

I will call tentative answers to scientific questions *hypotheses*, where a hypothesis can treat of some particular fact or event, or state a generalization. A *grammar* of L, being an answer to a scientific question (namely, Q) is a hypothesis; a hypothesis about L. (Or, if you prefer, a grammar of L is a related series of hypotheses about L.)

One tests a hypothesis against the *experiential data*, where this data seems to divide into two classes, for any hypothesis H: that data which is relevant to H (i.e. the *evidence* with respect to H) and that data which is irrelevant to H. That is:

(5)



Applied to grammars of particular languages: a piece of data D counts as relevant to G (i.e. evidence for or against G) if and only if (a) D confirms or disconfirms some part of G or (b) D confirms or disconfirms some hypothesis H, such that H confirms or disconfirms some part of G. My concern in this paper is: what data are relevant to testing grammars of particular languages?[1] Put otherwise: what counts as evidence for/against grammars of particular languages? My particular concern will be the evidentiary role, in linguistics, of texts — where by "texts" I mean *communicative events:* naturally occurring instances of language, whether spoken or written.[2]

2. Against the Strong Claim

If texts are the *only* data relevant to grammars, then they must provide all the data a linguist needs to answer Q. But, as a matter of fact, texts are inherently limited in the kinds of evidence they provide, in a number of important ways. In what follows I will consider only two kinds of examples. Many others could be added. The conclusion will be as follows: Since texts are not sufficient for constructing and testing grammars, they cannot be the only evidence in linguistics.

2.1 Grammaticality

In writing a grammar for a language L, one of the central tasks is to characterize each expression as fully grammatical, fully ungrammatical, or somewhere in between. For instance, any satisfactory grammar of English must have as a consequence that (6a) is well-formed, (6b) is slightly ill-formed, (6c) is more ill-formed, and (6d) is wholly ungrammatical.

(6) (a) The baby seems to be sleeping
(b) ?Seems the baby to be sleeping?
(c) *The baby seems sleeping
(d) **Seems the baby sleeping?

1. Let me stress: I do not think one can give an "in principle" answer to the question of what data are relevant to testing grammars. More on this below. On the other hand, one can already say with a fair degree of certainty what data have *as a matter of fact* been relevant up until now, and what data are likely to continue to be relevant.

2. My usage here follows Halliday (1961: 243) who writes: "The data to be accounted for are observed language events, observed as spoken or as codified in writing, any corpus of which, when used as material for linguistic description, is a 'text'."

Any grammar which (rightly) entails that (6a) is grammatically well-formed is nevertheless unsatisfactory if it does not also entail that (6b) through (6d) are grammatically ill-formed, to some degree.

So: the linguist, in writing her grammar, must garner evidence of grammaticality — and degrees thereof — from somewhere. But here is the rub: she cannot get this evidence from texts. The reason is simple: texts contain ungrammatical strings, and fail to contain all the grammatical ones. Furthermore, texts cannot establish the *degree* of grammatical well-formedness of the expressions they contain. Hence texts cannot be the only evidence in linguistics.

An example: undoubtedly, people can say "the baby seems sleeping" — and I wouldn't be surprised if someone has. Indeed, people can and do say all kinds of things: they produce language-like noises ("chuga, chuga"), they make false starts, they have heart attacks in mid-utterance, and so on. Sometimes they may even utter complete nonsense, just because they enjoy fooling linguists. Nevertheless, whether said or not, these things are not grammatical expressions of the language under study. In a word: being uttered is not a sufficient condition for being a grammatical expression.

Nor is being uttered a necessary condition for grammaticality. There are a whole host of grammatical expressions that have never been uttered. This is so even if one insists that languages contain only a finite number of expressions. After all, there will inevitably be *some* expressions which have never been uttered simply because no one has ever felt the urge to utter them.

This much has long been obvious. The response has sometimes been to introduce notions such as "likelihood of utterance." But to say that an expression is grammatical is not to say that it is likely to be uttered. For one thing, many fully grammatical expressions are extremely unlikely ever to be said; just as unlikely as wholly ungrammatical expressions. For another, some ungrammatical expressions have a non-negligible chance of being uttered; at least as high as some grammatical expressions. For example, compare (6c) and (7).

> (7) For the most part, laser printers are more expensive than toothpicks

"The baby seems sleeping" is clearly less grammatical than (7). Nevertheless I suspect that, statistically speaking, it is no less likely to occur.

Similarly, though (8) is slightly ill-formed, I suspect it is more likely to be uttered than the perfectly grammatical (7).

(8) *I taked it home

In sum: that a sentence appears in some text *does not* imply that it is grammatically well-formed. What's more, that a sentence fails to appear in any text *does not* imply that it is grammatically ill-formed. Still less do facts about texts determine the relative degree of grammaticality of the expressions contained therein. So texts do not provide one of the essential tests for and against grammars: evidence about grammaticality and degrees thereof.

One side note, before I go on. It is notoriously difficult to give a non-circular analysis of grammaticality. "Grammatical", it is often said, does not mean the same as "acceptable" or "usable" or what have you. What "grammatical" means is: generated by the grammar. This looks viciously circular. Personally, I'm optimistic about the possibilities of explicating grammaticality by reference to its place in linguistic theory as a whole. But, even if talk of "grammaticality" and "well-formedness" must ultimately be rejected, this in no way affects the arguments just presented. For the arguments can be modified to avoid reference to grammaticality. Notice:

Premise 1: Any adequate grammar must distinguish those expressions which 'sound right' from those which don't.

Premise 2: Very many expressions which 'sound right' have never been uttered.

Premise 3: Very many expressions which *don't* 'sound right' have been uttered.

Premise 4: There is no sufficiently reliable statistical correlation between 'sounding right' and being uttered.

Therefore,

Conclusion 1: Texts cannot distinguish expressions which 'sound right' from expressions which don't.

Conclusion 2: Texts do not provide all the evidence a linguist needs to construct a grammar for L and test it.

2.2 Interpretability

Being grammatical and being interpretable are not the same thing. A sentence may be perfectly grammatical, but very hard to interpret. Sentences (9) and (10) are familiar examples.

(9) The horse paraded past the stand died
(10) The hay the horse the man rode ate was very dry

Furthermore, a sentence may be ungrammatical, but easy to interpret. Though (6b) is slightly ill-formed, it is fairly easy to understand. The same holds for (6c).

A linguist should, it must be agreed, use data about interpretability — and not just facts about grammaticality — when evaluating her grammar. For example: a grammar of English which had as a consequence that (6b) conveyed no meaning whatever would be deficient; as would a grammar that had as a consequence that (9) and (10) are nonsense — rather than simply being difficult to understand.

But how can texts provide evidence from interpretability? Once again, that an expression fails to appear *does not* imply that it is uninterpretable. Lots of interpretable expressions are never uttered: because they're slightly ill-formed, because they're too hard to understand, or just because no one has ever felt an urge to say them. And, as is depressingly familiar to those who read undergraduate essays, being produced is no guarantee of being interpretable.

The source of the problem is this: to say that an expression E is interpretable is to say that someone who knows the language, and has the other requisite mental abilities, *is able to understand* E. But, evidently, human beings haven't actually said everything that they can understand. Nor is everything that humans say inevitably something which you, I, or anyone else, can understand.

Of course once talk of 'understanding' enters in, the problem with the Strong Claim becomes clear. In testing grammars, reference to mentalistic notions — including native speaker's intuitions about what 'sounds right' and what can be understood — is ubiquitous. And if the failure of behaviourism has one lesson, it's this: you can't operationally define mentalistic notions using purely behavioural language. No amount of talk about 'dispositions to behave' is enough to cash out understanding, believing, wanting and so on. Still less can you cash out understanding and such in terms of *some subset of actual past behaviour* — e.g. texts.

To sum up: so far as I can tell no amount of text can ever suffice to categorize expressions in terms of their degree of grammaticality or their degree of interpretability. I could go on with further examples: whether two expressions are synonymous ("hazel nut" and "filbert"); whether one

entails the other ("John persuaded Sue to leave" entails "Sue decided to leave"); whether an expression is ambiguous; whether an expression is a possible (versus impossible) word in L (e.g. "zoar" is possible in English, but "ptlump" is not); and so on. None of this evidence about particular languages can be garnered from texts, so far as I can see. Hence texts cannot be the only evidence in linguistics.

2.3 The Language Acquisition Reply

There is a natural defense of the Strong Claim that goes like this. Children construct a grammar of the language spoken around them. *They* don't elicit speaker's judgements, nor do they have the judgements themselves — the only information available to children are actual texts. Therefore, it must be possible to construct a grammar from texts alone, because children do it.

This argument is valid, but it contains a false premise: that the only information available to children are actual texts. There is every reason to think that children have a "head start" in grammar construction — i.e. they have an innate linguistic endowment — which linguists do not have access to. Children can make do with much less evidence than linguists can because they start off knowing much more than linguists do. So it just isn't true that the only information available to children about their language comes through texts. That is why, though intuitive judgements aren't required for *language acquisition*, they are necessary in *linguistics*.

2.4 The Ought vs Is Reply

To the claim that texts are not the only evidence in linguistics it might reasonably be replied that texts *should be* the only evidence in linguistics. That is, if practising linguists use other kinds of evidence, they are wrong to do so. This is precisely the line that methodological behaviourists have taken in psychology. I think it's a mistake.

First of all, restricting oneself to texts would amount to ruling out, strictly a priori, a whole range of data. No other natural science does this. Hence linguistics shouldn't. And too: there is an awfully good reason why other sciences do not rule out any evidence a priori. The reason is, confirmation in science is *isotropic*. What is isotropy? Fodor (1983: 105) explains it as follows:

> By saying that confirmation is isotropic, I mean that the facts relevant to the confirmation of a scientific hypothesis may be drawn from anywhere in the field of previously established empirical (or, of course, demonstrative) truths. Crudely: everything that the scientist knows is, in principle, relevant to determining what else he ought to believe. In principle, our botany constrains our astronomy, if only we could think of ways to make them connect.

The fact is, nobody currently knows how things are connected up. Hence one shouldn't assume, a priori, that certain connections *could not hold*. That's why it may turn out, for example, that the moon's phases affect plant growth. And that's why confirmation in science is (and should be) isotropic.[1]

Carl Hempel (1966: 12) makes another important point:

> . . . what particular sort of data it is reasonable to collect is not determined by the problem under study, but by a tentative answer to it that the investigator entertains in the form of a conjecture or hypothesis.

Hempel offers the case of Ignaz Semmelweis as an illustration. Semmelweis discovered that a large proportion of the women who delivered their babies in the First Division of the Vienna General Hospital contracted a serious and often fatal illness known as 'childbed fever'. In the adjacent Second Maternity Division of the same hospital, which accommodated almost as many women as the First, the death toll from childbed fever was much lower (Hempel 1966: 3). The interesting thing about this case, for my purposes, is an explanation of these facts which Semmelweis considered and rejected, thereby making progress on the problem. It was pointed out that, because of the floor plan of the First Division,

---

1. Scientific reasoning is isotropic in another sense too: two processes which seem to be mutually irrelevant may illuminate each other *by analogy*. For example, what we know about water flow gets borrowed to model the flow of electricity; and what we hypothesize about the structure of the solar system is used to model the structure of the atom. Here too, then, everything that the scientist knows is, in principle, relevant to determining what else she ought to believe — because analogies can occur in the craziest of places. Hence it's just plain foolish to refuse to look in certain places for evidence. (It's worth stressing, however, that analogies are not necessarily *evidence* with respect to the hypotheses they give rise to. An analogy might well help one formulate a hypothesis, without confirming or disconfirming it.)

> . . . a priest bearing the last sacrament to a dying woman had to pass through five wards before reaching the sickroom beyond: the appearance of the priest, preceded by an attendant ringing a bell, was held to have a terrifying and debilitating effect upon the patients in the wards and thus to make them more likely victims of childbed fever. In the Second Division, this adverse factor was absent, since the priest had direct access to the sickroom. (Hempel 1966: 4)

Semmelweis tested this conjecture by nixing the bell and having the priest take an indirect route to the sickroom. No difference in mortality was observed. Hence the hypothesis was disconfirmed.

This example shows that data are relevant or irrelevant only *relative to a hypothesis being tested*. There is no such thing as '*the* data relevant to an enterprise'. In this instance, data about the priest's walking behaviour turned out, oddly enough, to be relevant to the problem of why childbed fever was more common in the First Division of Vienna General Hospital in the 1840s. Similar bizarre happenings can be expected in linguistics. The lesson: there is no saying, in principle, what data will be relevant in linguistics. That will necessarily depend on which hypotheses linguists form and wish to test. Hence it would be foolhardy to rule out any evidence a priori.

So the Strong Claim is wrong. Texts are not, and should not be, the only acceptable evidence in linguistics.

3. Against The Weaker Claim

I turn now to the Weaker Claim, repeated below:

> (2) *The Weaker Claim:* Texts are the best evidence in linguistics; (better, in any case, than native speaker's intuitions)

The thought here is that it is better to have a text as evidence, when this is possible. I disagree, for two reasons. First, texts are not *more reliable* evidence. Second, texts are less '*cost effective*' evidence — hence they are less relevant, since relevance is a function of degree of confirmation/disconfirmation *and cost*. (See Sperber and Wilson 1986.)

3.1 Texts are 'Messy'

As Chomsky has repeatedly stressed, "The actual use of language involves a complex interplay of many factors of the most disparate sort, of which grammatical processes constitute only one" (Chomsky 1964: 54). As a result, texts tend to be extremely 'messy'. To separate the irrelevant data from the relevant data (i.e. the evidence), therefore, a linguist must decide which features of the messy result derive from the language being used, and which derive from other factors. To put this another way: she must decide which behaviours are truly *linguistic behaviours*.

In practice, the selectivity imposed by the 'messiness' of text is still more acute. For, beyond deciding what is linguistic behaviour and what is not, a linguist wishing to appeal to texts must:

1. decide which speech episodes to record;
2. decide which recordings to transcribe;
3. decide which transcribed texts to use.

Allow me to give a personal example. Several years ago, Alisa Hillier and I edited a volume of texts called *A LE CAMP Sourcebook*. (Stainton and Hillier 1990. A small excerpt of this volume appears in the Birthday Text.) This *Sourcebook* consists of transcriptions of video tapes, taken at a French immersion camp. The first two years, our team made a dozen or so tapes — 3 hours each. That's about 36 hours of tape. This represented four camp sessions, six weeks each, 30 hours per week. Some quick math: that comes to 720 hours that the children were on site — of which only 36 hours was recorded! The tapes were transcribed by a team of corpus-based linguists in Israel and in Canada. From nearly 40 hours of recordings, Hillier and I produced the *Sourcebook* — which contains only 243 pages of text. And our reports, based on the LE CAMP research, made use of something like twenty percent of the book: roughly 50 pages. So: 720 hours of 'linguistic behaviour' translates into 50 pages of useful text. Why so little? Because we — myself included — judiciously selected things to film, chose appropriate bits of film to transcribe, and picked as our evidence portions of the transcript which were clean enough to be useful. Finally, having selected a text, we abstracted away from those features of it which were irrelevant (e.g. noise, false starts, and so on).

Let me stress: this selectivity was not a *failing*, peculiar to our research. To make corpus-based linguistics manageable at all, one simply must be selective. Now, consider: what do text-based linguists rely upon when 'cleaning up' a corpus? The answer, from personal experience, is clear: they

use their intuitions as speakers of the language. But if the use of texts *presupposes* appeal to native speaker intuitions, the reliability of the former can never exceed that of the latter. Hence texts are not better (read: more reliable) evidence than native speaker's judgements.

3.2 The 'Cost Effectiveness' of Texts

The claim under consideration is that it is better to have a text as evidence, when this is possible. This claim is surely incorrect, when understood such that a linguist should seek out a text as evidence whenever she can. The reason is that, even where it is *possible* to find or record a text, doing so is not always the best course of action; it may be a waste of valuable time and energy. This will happen whenever there is an easier way of confirming or disconfirming a hypothesis.

Here is an example. Suppose Andrea is working on a grammar of French. She hypothesizes that French is not a pro-drop language. (That is, NPs lacking a phonetic 'spell out' cannot occur in subject position of finite clauses.) She wishes to confirm this hypothesis. But how? She could do a statistical analysis of thousands of hours of speech, to show that subjectless sentences very rarely occur. She might even hope to find a text in which an utterance of a subjectless sentence prompts what Quine (1953: 53) calls 'a bizarreness reaction'. This wouldn't *establish* that French is not a pro-drop language: as I argued above, the absence of an expression does not show it to be ungrammatical, or uninterpretable. Nor does a bizarreness reaction show that the expression uttered is ungrammatical, or uninterpretable. (E.g. speakers might refrain from using phonetically null elements in subject position because this is taboo; this would explain the bizarreness reaction. See Stich 1972 for a useful discussion.) But, for the sake of argument, suppose that these thousands of pages of text partially confirm Andrea's hypothesis. The question is: ought Andrea to pursue *this* line of research to confirm her hypothesis? Only someone in the grip of an ideology could think so. What Andrea should do is present native French speakers with subjectless sentences, and ask them if they 'sound right'. (And, if Andrea speaks French herself, she should use her own judgements as well.)

In this example, it clearly would *not* be better to make use of texts — even though this is possible. That is because gathering and analyzing texts is not the best (read: most cost efficient) means of verifying the hypothesis in question. And there is a good reason for this: in linguistics, as in science

generally, nature doesn't usually tell the investigator what she wants to know; the investigator has to ask. (Compare chemistry. Imagine a chemist wishing to verify her prediction that chemicals A and B, when combined in mixture M, will explode. Would she be wise to wait around and hope for chemicals to combine in just the right proportion? Clearly not. What she ought to do is measure out the chemicals, put them together, and stand back.)

In a word: straight observation is a rather poor methodology in science. Experimentation is usually the best way to go — in linguistics, and in science generally. As Hempel (1966: 17) writes:

> Scientific knowledge . . . is not arrived at by applying some inductive inference procedure to antecedently collected data, but rather by what is often called "the method of hypothesis", i.e. by inventing hypotheses as tentative answers to a problem under study, and then subjecting these to empirical test.

3.3 Reply: Intuitions are Unreliable

The Weaker Claim might be defended in the following way. Granted, there are problems with texts. They are messy. And it is difficult to collect appropriate ones. But — the objection goes — texts are nevertheless better than speaker's intuitions, because the latter are *so* unreliable.

Here is a famous example of such unreliability. In his widely read polemic against what he called "transformational analysts", Archibald Hill (1961) concluded that speakers cannot distinguish well-formed from ill-formed sentences in their language. It follows that speaker's judgements of grammaticality are completely unreliable. (For discussion, see Newmeyer 1983: 63ff.)

Hill based his conclusion on an experiment, in which subjects were presented cards on which were written the following example sentences: ([11a] through [11j] were taken from Chomsky's *Syntactic Structures*).

(11) (a) Colorless green ideas sleep furiously
(b) Furiously sleep ideas green colorless
(c) have you a book on modern music?
(d) the book seems interesting
(e) read you a book on modern music?
(f) the child seems sleeping

(g) I saw a fragile whale
(h) I saw a fragile of
(i) Those man left yesterday
(j) I never heard a green horse smoke a dozen oranges

Hill's informants were asked, "to reject any sentences which were ungrammatical, and to accept those which were grammatical." Not to put the point too finely, the results were pathetic. I cite Hill's report in full.

> In the results of this voting, (11i) 'those man . . .' was the only sentence rejected by all ten informants. Sentence (11b) 'Furiously sleep . . .' was rejected by seven, accepted by three. The voting on (11h) 'fragile of.' was similar, seven rejecting, three (but not the same three) accepting. Sentences (11e) 'read you . . . ?' and (11f) 'seems sleeping.' were each rejected by four, accepted by six. Sentence (11j) '. . . green horse' was rejected by one, accepted by nine. Sentence (11c) 'have you a book . . .' was accepted by all but one informant, who offered the qualification that it would be ungrammatical in his idiolect. This informant was one of the two linguists fully aware of differences in British and American dialects. Sentence (11d) 'the book seems . . .' was hesitated over before final acceptance by one informant. All others accepted without hesitation. Sentence (11g) '. . . fragile whale,' was accepted without hesitation by all informants. (Hill 1961: 3. The numbering system is mine.)

What do these results suggest? Well, certainly one can say this: the informants' judgements did not *support* Chomsky's claims about which of (11a) through (11j) were well-formed. But Hill draws a stronger conclusion. He maintains that this amount of disagreement between informants establishes that native speaker's judgements of grammaticality, when elicited under these sorts of conditions, simply are not the least bit reliable.[1]

1. A similar experiment was conducted by Spencer, with predictably analogous results. Spencer (1973: 83) writes: "One hundred and fifty exemplar sentences from 6 linguists' articles were presented to 43 linguistically naive and 22 linguistically nonnaive speakers. Native speakers agreed among themselves as to the acceptability or unacceptability of 80% of the sentences. Subjects shared intuitions with linguists in only half of the exemplars."

How to respond? Does Hill's experiment really show that intuitions are wholly unreliable — thus saving the Weaker Claim? I think not. In the first place, no practising linguist would elicit speakers' judgements *under these sorts of conditions*. To get an informative answer, one must adequately explain one's question. Applied to this case: the linguist must explain to the informants that she doesn't want to know whether the sentences are *interpretable*, or *easy to parse*, or whether they could *occur as a line of poetry*. Explaining the query would undoubtedly improve the informants' performance.

But still: even explaining — as best one can — what is meant by "grammatical" would not solve the problem altogether. Anyone who has ever tested a hypothesis on untutored speakers knows that, no matter how carefully one explains the question, answers can vary between speakers of the very same dialect. Furthermore, as Labov (1975) has rightly stressed, linguists' judgements tend to be affected by their stance on an issue, not to mention intuitions being clouded by over exposure and such. In short, there's no denying that speakers' reports — whether from untutored informants or from trained linguists — are not wholly reliable.

But why is *that* a problem? *No* data in *any* scientific endeavour is 100% trustworthy. Gauges malfunction, scientists observe what they hope to find, experiments don't turn out, test subjects make mistakes, and so on — in *every* empirical enterprise. If *this* were enough to discredit a discipline, the whole of science would be in jeopardy. Don't get me wrong: there is a real problem here. But, I repeat, it's a problem for linguistics and for every other empirical science. There is, I fear, no in principle solution. All one can do is exercise the appropriate level of caution: re-check results, make them available for independent test, seek out clear cases, use data from different sources, and so on. What one should not do is simply discard all evidence that is not completely reliable. That wouldn't be scientific caution; that would be scientific suicide.

Besides, the issue was never whether intuitive judgements were 100% trustworthy as evidence. Of course they're not.[1] Rather, in evaluating the Weaker Claim, the issue was and is: which is more difficult to obtain, reliable and relevant *reports*, or reliable and relevant *texts*? The answer is, reliable texts are harder to come by because: (a) one must use native speaker's intuitions to select and edit the texts in the first place; and (b) one can

1. Nor was Chomsky under any illusions about this. As he wrote in reply to Hill, way back in 1961: ." . . the difficulty of obtaining reliable and relevant reports is quite apparent" (Chomsky 1961: 225).

reduce error in reports by framing the question appropriately, but there's no obvious means of reducing error in texts. Furthermore, one can compare the judgements of untutored informants with those of linguists, and vice versa, thereby (hopefully) finding the underlying pattern in the surface mess.

4. The Weakest Claim

The preceding discussion, in which I try to establish that both the Strong Claim and the Weaker Claim are mistaken, should be sufficient to cast doubt on the Weakest Claim as well. It is not the least bit clear that texts are very good evidence: because they are too messy, and require too much work; and because, in any case, they are incapable of assigning degrees of grammaticality and interpretability.

But are texts any good at all? Of course they are. Let me end my discussion by considering some of the possible merits of texts.

4.1 Rejecting/Clarifying Speaker's Judgements

As I have said, an informant's reports of linguistic intuitions are not wholly reliable. For one thing, intuitions — like texts — result from an interplay of causes. The result is that, for example, (sometimes) informants deny that a perfectly well-formed expression is grammatical. Another example: informants may insist that a sentence which is, in fact, interpretable — i.e. is assigned a meaning by the grammar of the language — cannot be understood. (Sentences [9] and [10] provide obvious candidates for being incorrectly labelled ungrammatical and/or uninterpretable.)

Suppose a linguist encounters such a case, in which she suspects that her informants are mistaken about the grammaticality or interpretability of an expression in their language. Suppose further that, as it turns out, these informants regularly utter this very expression — thereby producing many texts which contain it. This is one case in which texts are useful — they allow the linguist to discount an intuitive judgement as unreliable. As Stich (1972: 130) says:

> An informant's protest that a given sequence is unacceptable may be ignored if he is caught in the act, regularly uttering unpremeditatedly what, on meditation, he alleges he doesn't say.

Again: in constructing grammars, one often encounters expressions whose grammaticality/interpretability is in doubt because of disagreements between informants. Different informants sometimes disagree about the degree of well-formedness/interpretability of a given expression. Here too, the regular occurrence of such expressions in actual speech may rightly cause a linguist to classify the in-doubt expressions as grammatical/interpretable, despite sincere avowals to the contrary by certain informants.

So, texts are evidence in linguistics. But notice what these two cases have in common: it is only when the informants' judgements are questionable that one makes recourse to texts. Texts play a role when intuitions conflict with one's predictions; or when intuitions are rather unclear. In clear cut cases, texts need not — and hence should not — be sought out.[1]

5. Conclusion

By way of conclusion, let me be clear about what I am saying, and what I'm not. I am not saying that, in principle, texts cannot be relevant to linguistic hypotheses. I don't think anything can be ruled out as linguistic evidence — in principle. Nor do I believe that, in practice, texts are not a source of evidence. I have been concerned to argue that texts have no pride of place in confirming grammars (i.e. the Strong Claim and the Weaker Claim are mistaken); and — most important of all — that the demand for 'real language' in the form of a text is generally unwarranted and unproductive.

1. The most useful feature of texts *for me* has been as a source of ideas. My work on minor clauses was prompted by looking at the aforementioned LE CAMP texts: I noted how exceedingly common they were, and how easily understood. And I realized that very little had been written about the syntax, semantics and pragmatics of minor clauses. It may be that I would never have pursued this topic, but for looking closely at texts. And, since texts include both recorded speech episodes plus bus posters, novels, letters and the like, I can safely say that texts have drawn my attention to any number of facts about English.

Then again, a source of ideas isn't really the same thing as *evidence*. (See my remarks about analogies.) Staring at a photograph, or watching a ballet, can as easily produce a brainstorm as a poster. But they don't necessarily form part of the evidence for the resulting hypothesis.

Another place where texts play an important role is in the study of conversation: exchange structures, discourse schemata, genres and so forth. At first glance, it would seem that the only viable way to find these complex patterns, distributed over lengthy discourse chunks, is to examine texts. (I owe this point to William Downes, in conversation.)

References

CHOMSKY, N. (1957). *Syntactic Structures*. 'S-Gravenhage: Mouton and Co.

— (1961). "Some methodological remarks on generative grammar." *Word* 17:219-239.

— (1964). "Current issues in linguistic theory." In J.A. FODOR and J.J. KATZ, eds. *The Structure of Language. Readings in the Philosophy of Language*. Englewood Cliffs, N.J.: Prentice-Hall.

FODOR, J.A. (1983). *The Modularity of Mind*. Cambridge, MA: The MIT Press.

GOODMAN, N. (1954). *Fact, Fiction and Forecast*. Cambridge, MA: Harvard University Press.

HALLIDAY, M.A.K. (1961). "Categories of the theory of grammar." *Word* 17:241-292.

HEMPEL, C.G. (1966). *Philosophy of Natural Science*. Englewood Cliffs, N.J.: Prentice-Hall.

HILL, A.A. (1961). "Grammaticality." *Word* 17:1-10.

KRIPKE, S. (1982). "On rules and private language." Reprinted in A.P. MARTINICH, ed. (1990). *The Philosophy of Language*. 2nd ed. Oxford: Oxford University Press.

LABOV, W. (1975). *What is a Linguistic Fact?* Lisse: Peter de Ridder Press.

NEWMEYER, F.J. (1983). *Grammatical Theory: Its Limits and Its Possibilities*. Chicago: The University of Chicago Press.

QUINE, W.V.O. (1960). *Word and Object*. Cambridge, MA: The MIT Press.

— (1953). "The problem of meaning in linguistics." Reprinted in J.J. KATZ, ed. (1985). *The Philosophy of Linguistics*. Oxford: Oxford University Press.

SPENCER, N.J. (1973). "Differences between linguists and nonlinguists in intuitions of grammaticality-acceptability." *Journal of Psycholinguistic Research* 2:83-98.

SPERBER, D. and D. WILSON (1986). *Relevance*. Cambridge, MA: Harvard University Press.

STAINTON, R.J. and A. HILLIER, eds. (1990). *A LE CAMP Sourcebook* Vol. 1. Toronto: Applied Linguistics Research Working Group.

STICH, S. (1972). "Grammar, psychology and indeterminacy." Reprinted in J.J. KATZ, ed. (1985). *The Philosophy of Linguistics*. Oxford: Oxford University Press.