

# The Legacy of Methodological Dualism\*

Word Count: 15202

ABSTRACT. Methodological dualism in linguistics occurs when its theories are subjected to standards that are inappropriate for them qua scientific theories. Despite much opposition, methodological dualism abounds in contemporary thinking. In this paper, I treat linguistics as a scientific activity and explore some instances of dualism. By extracting some ubiquitous aspects of scientific methodology from its typically quantitative expression, I show that two recent instances of methodologically dualistic critiques of linguistics are ill-founded. I then show that there are nonetheless some divergences between linguistic and other ordinary scientific methods, reflecting yet a third instance of methodological dualism.

## 0 Introduction

---

I am grateful to Jeff Barrett, Hans Kamp, Kyle Stanford, the editors of this journal, and two anonymous reviewers for reading complete drafts of this paper and providing much useful feedback. Versions of this paper were presented at the University of Michigan Linguistics and Philosophy Workshop and at the Cal State Long Beach conference on The Epistemology of Natural and Artificial Systems. Several audience members provided either helpful comments or interesting sociological data points.

**Address for correspondence:** Department of Logic and Philosophy of Science, 3151 SSPA, University of California, Irvine, Irvine, CA 92697-5100.

**Email:** johnsonk@uci.edu.

Perhaps more than any other discipline, linguistics has continually defended its methods and practices as ‘scientific’. This practice was heavily inspired by Noam Chomsky’s frequent and vigorous critiques of ‘methodological dualism’. Methodological dualism occurs when a discipline’s theories are subjected to standards which are inappropriate for them qua scientific theories. In this vein, Chomsky repeatedly charged many central figures in philosophy – Dummett, Davidson, Kripke, Putnam, and Quine, just to name a few – of subjecting theories of language (and mind) to dualistic standards (e.g., Chomsky 1986, ch. 4; 2000, ch. 2 – 6; 1975, ch. 4). Moreover, he holds that these standards have no known plausible defense, and that there is no reason to take them seriously. In place of these dualistic requirements, Chomsky recommends that linguistic theories be held to the standards that normally apply to empirical theories. In his words, he advocates ‘an approach to the mind that considers language and similar phenomena to be elements of the natural world, to be studied by ordinary methods of empirical inquiry’ (Chomsky 2000, 106). This is a natural position for Chomsky, given his rather traditional view of the relationship between philosophy and science:

In discussing the intellectual tradition in which I believe contemporary work [sc. on language] finds its natural place, I do not make a sharp distinction between philosophy and science. The distinction, justifiable or not, is a fairly recent one.... What we call [Descartes’] ‘philosophical work’ is not separable from his ‘scientific work’ but is rather a component of it concerned with the conceptual foundations of science and the outer reaches of scientific speculation and (in his eyes) inference. (Chomsky 1988, 2).

If philosophy is a kind of study into the foundations of the sciences, then there is little room for a ‘philosophical’ theory of language or mind that is not itself a ‘scientific’ theory.

The view that linguistic theories are ordinary scientific theories, subject to the same methodological standards as the (other) sciences, has been endorsed, at least in name, by

virtually all linguists and a great many philosophers. The aim of this paper is to explore this view and its concomitant rejection of methodological dualism.<sup>1</sup> A bit more specifically, I take three cases of alleged methodological dualism, and use them to consider the general question of how linguistics fares with respect to the ordinary standards and methods that pervade virtually all of the empirical sciences. The results, I argue, are mixed. The first two cases (discussed in §§1 – 2), were put forward by prominent philosophers and linguists. They argue that there are deep and principled difficulties with some common practices in contemporary linguistics. But in both cases, we will uncover a remarkably tight point-by-point agreement between the relevant aspects of linguistic methods and the underlying logic of the other sciences. Thus, if linguistics is subject to the same standards as the other sciences, and if (as I assume) the other sciences are not extremely vitiated – probably beyond repair – then both charges must be rejected as cases of unwarranted methodological dualism.

The third case (discussed in §3) is different. It focuses on the rather striking absence of quantitative methods in linguistic theorizing. It is no accident that science is based on numbers and numerical algorithms. The numerical representation of data and the mathematical manipulation of these representations allow for precise solutions that humans are notoriously bad

---

<sup>1</sup> One occasionally hears that linguistic theories needn't be treated as scientific, but can be thought of as 'philosophical' instead. Whatever this position amounts to, several things can be said in response. First, such alternative theories don't appear to compete or conflict with anything I say about scientific theories. Second, I mean very little by calling a semantic theory 'scientific'. Linguistic theories may be considered 'scientific' primarily because their construction and confirmation centrally involve employing some of our best-known methods for obtaining knowledge about a particular empirical phenomenon. From this perspective, it is unclear what the *value* of a 'non-scientific' theory of language would be. Moreover, my use of the assumption that linguistic theories are a type of scientific theory is especially uncontentious. So for an alternative view to avoid my conclusions, one needs to show why the particular features of linguistic theorizing that I employ are not part of some other *worthwhile* form of linguistics.

at estimating with subjective judgments. The topic of §3 presents an exciting and promising area of research into linguistics and its methodology, and is pursued in greater depth in Johnson ms.. Here, though, here I focus solely on identifying the relevant issue. I conclude in §4.

I should note that my use of ‘linguistics’ in this paper is limited to research within the tradition of generative grammar, broadly conceived, such as the research inspired by Chomsky’s work (e.g., 1965, 1995). I have little to say about the many other rich and important projects that also take place in linguistics departments.

## **1 Dualism I: Saving the Phenomena**

The first form of methodological dualism I discuss concerns the explanatory scope of a semantic theory. Francois Recanati and others maintain that semantic theories must capture a great deal of the apparent semantic phenomena, in a sense to be spelled out below. This requirement is used to critique various types of semantic theories, such as the one developed by Herman Cappelen and Ernie Lepore (hereafter CL). I briefly sketch CL’s view, and then consider Recanati’s criticism. We’ll then be able to see the dualistic nature behind the requirement that semantic theories ‘save the phenomena’.

### **1.1 The Proposed Linguistic Theory**

CL recently developed a view called ‘Semantic Minimalism’ (CL 2005). According to Semantic Minimalism, only a handful of expressions are actually context sensitive (e.g., *I*, *you*, *she*, *this*, *that*, *tomorrow*, etc.). There are no hidden (i.e., unpronounced, unwritten) context-sensitive elements in the syntactic or semantic structure of an expression. Thus, Semantic Minimalism

contrasts sharply with most semantic theories. In particular, it is normal for semanticists and philosophers of language to assume that a correct semantics for (1a) – (2a) is something along the lines of (1b) – (2b), where more structure is assigned to the semantics than what is given by the overt syntactic structure of these sentences.

- (1) a. Mary is ready;  
       b. Mary is ready *to X*.
- (2) a. It is raining;  
       b. It is raining *in location X*.

Thus, (1a) means something like Mary is ready to do some particular salient activity, or is ready for something in particular to happen. Similarly, (2a) means it's raining in some contextually specified place. Semantic Minimalism denies this, holding instead that (1a) simply means that Mary is ready, and (2a) simply means that it's raining.

## 1.2 Criticism of the Theory via Methodological Principle

Unsurprisingly, Semantic Minimalism has encountered numerous objections (cf. CL 2005, ch. 11 – 12 for discussion). I focus on just one of these, which I call the ‘Problem of Unsaved Phenomena’ (PUP). PUP is most straightforwardly presented in Recanati 2001 (cf. also Carston 2004 Allen 2003 (552) for similar sentiments, and CL 2005, ch. 12 and citations therein). For instance, Recanati writes:

That minimal notion of what is said is an abstraction with no psychological reality, because of the holistic nature of speaker's meaning. From a psychological point of view, we cannot separate those aspects of speaker's meaning which fills gaps in the representation associated with the sentence as a result of purely semantic interpretation, and those aspects of speaker's meaning which are optional and enrich or otherwise modify the representation in question. They are indissociable, mutually dependent aspects of a single process of pragmatic interpretation. (Recanati 2001, 88)

Recanati's pessimism about Minimalist semantic theories is driven largely by his view that such theories don't explain enough of the phenomena. In Recanati's view, a semantic theory must capture the (entire) 'content of the statement as the participants in the conversation themselves would gloss it' (Recanati 2001, 79-80). Recanati expresses this in his

'Availability Principle', according to which 'what is said' must be analyzed in conformity to the intuitions shared by those who fully understand the utterance – typically the speaker and the hearer, in a normal conversational setting. This in turn supports the claim that the optional elements... (e.g., the reference to a particular time in 'I've had breakfast') are indeed constitutive of what is said, despite their optional character. For if we subtract those elements, the resulting proposition no longer corresponds to the intuitive truth conditions of the utterance. (Recanati 2001, 80)

Prima facie, Recanati's attitude seems quite reasonable: the least we can demand of a theory is that it 'save the phenomena'. Only after we have theories that do that should we consider their respective degrees of simplicity, elegance, etc. On further inspection, though, a different picture emerges. Let's begin by putting Recanati's argument in a more manageable form. We can characterize the Availability Principle as:

- (AP) A semantic theory is acceptable only if it correctly characterizes the intuitive truth conditions often enough within some psychologically interesting range of cases.<sup>2</sup>

Moreover, we can give PUP the following form:

*The Problem of Unsaved Phenomena*

- (3i) In all relevant ranges of cases, the intuitive truth conditions of our utterances contain much more content than what is characterized by minimalist theories.
- (3ii) If (3i) is right, then from a psychological point of view, we cannot separate the minimalist aspects of meaning from those aspects supplied by a more enriched view of meaning (often enough, in any relevant range of cases).
- (3iii) Hence, minimalist aspects of meaning cannot be separated from those aspects supplied by a more enriched view of meaning (often enough, in any relevant range of cases).
- (3iv) But if we can't separate minimalist from non-minimalist elements of meaning (often enough, in any relevant range of cases), then minimalist theories are unacceptable.
- (3v) Hence, minimalist views are unacceptable.

---

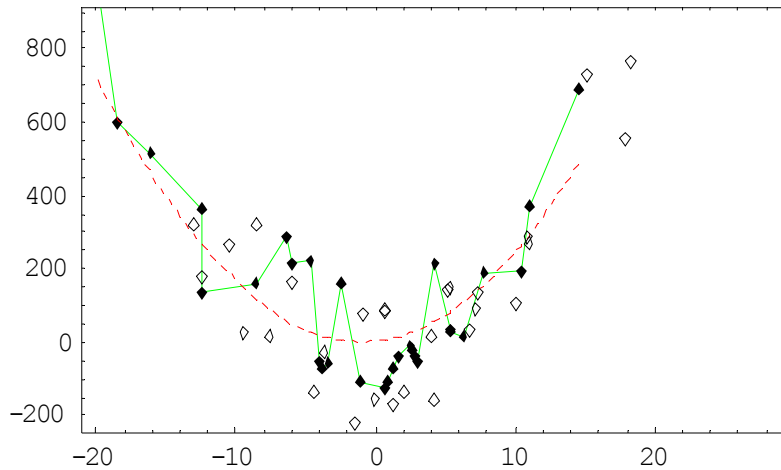
<sup>2</sup> For present purposes, I will assume that (AP) is an appropriate formulation of Recanati's Availability Principle; any divergences between the two will not matter in this paper. One might strengthen (AP) further by specifying the particular range of cases in which a semantic theory must get things right, and by specifying how often the theory must get things right. I won't worry about such strengthenings, though; since what I have to say will apply equally to all such versions of (AP).

Premise (3i) is an empirical claim; premises (3ii) and (3iv) are theoretical. Premise (3ii) comes from the quote of Recanati above (2001, 88), and premise (3iv) comes from (AP). (To see this, notice that if we can never separate out the minimalist aspects of meaning, then there must always be some non-minimalist aspects present, so the minimalist aspects of meaning never characterize the intuitive truth conditions in the utterance. Hence, by (AP), minimalist theories are unacceptable.) I won't discuss CL's attempt to deny the empirical claim (3i); suffice it to say that I find it inconclusive. But there are also flaws with (AP), (3ii) and (3iv), assuming linguistic theories are treated like ordinary scientific theories.

### **1.3 Is the Principle Justified by General Scientific Considerations?**

In order to see what is wrong with (AP), (3ii), and (3iv), it will be useful to step back from linguistic theorizing and examine some aspects of the methodology of the (other) sciences. I claim that non-linguistic scientific theories needn't observe appropriate counterparts of these principles. The parallel between linguistic and other scientific theories thus renders (AP), (3ii) and (3iv) unacceptable.

To get things started, let's take a simple example. Suppose we are studying the relationship between different quantities of a given additive X used in some manufacturing process and the amount of some type of atmospheric pollution Y generated by the process. The industry standard is to use  $n$  units of X per ton of product, but for a period of time, certain companies used more or less than  $n$  units. The relation between the varying amounts of X used and Y emitted are given as black diamonds in the plot below (ignore the two curves and white diamonds for the moment; also zero on the x-axis represents the use of  $n$  units of X; other values represent the respective deviations from this standard).



(The example of a pollution study here is arbitrary; it could be replaced with literally thousands of different examples from any given area of empirical science.) Given this data, there are infinitely many possible relations that could hold between  $X$  and  $Y$ . One extreme option would be to insist that every aspect of the data is crucial to understanding how  $X$  and  $Y$  are related. In such a case, a researcher might look for a function that captured the data precisely, as in the very complex one depicted with a solid line. In the present case, a polynomial of order 29, will do so, for the given raw data set of size 30:

$$(4) \quad \text{Predicted value of } Y_i = Y_i = f_2(X_i) = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \dots + \beta_{29} X_i^{29}$$

The resulting theory will then perfectly predict the behavior of  $Y$  on the basis of the behavior of  $X$ . The raw data, in the form of a set of pairs of measurements  $\{\langle X_i, Y_i \rangle : i \in I\}$ , is fully accounted for. In other words, (4) saves all the phenomena, which in this case is the variation in  $\langle X, Y \rangle$  scores of individual samples.

Despite its success at capturing the data, the first approach is almost never adopted. A vastly more common strategy hypothesizes a simpler relation between  $X$  and  $Y$ , and also that  $Y$  is influenced by other factors unrelated to  $X$ . One might, e.g., hypothesize that that relationship is given by the simple function:

$$(5) \quad \text{Predicted value of } Y_i = f_1(X_i) = \beta_0 + \beta_1 X_i + \beta_2 X_i^2$$

for some fixed numbers  $\beta_0, \beta_1, \beta_2$ . Once these numbers are determined from the data, we get the simpler curve given by the dashed line. In the present example, the values of  $\beta_0, \beta_1$ , and  $\beta_2$  were determined by seeking those values for which  $\sum_{i \in I} [(Y_i - f_1(X_i))^2]$  is as small as possible.

Although (5) doesn't predict the behavior of the original data as well as its rival (4), many other theoretical considerations speak in its favor. For example, suppose we got hold of another sample of data, given by the white diamonds above. Then we might ask how well the two functions captured this new data. One way to do this would be to compare the sizes of the discrepancies between what (5) and (4) predict about the value of  $Y$  for given values of  $X$  in the new data set. E.g., we might examine the ratio:

$$(6) \quad \frac{\sum_{i \in I'} [(Y_i - f_1(X_i))^2]}{\sum_{i \in I'} [(Y_i - f_2(X_i))^2]}$$

Here  $I'$  indexes the second set of measurements, and  $f_1$  and  $f_2$  are assumed to have had the particular numerical values of their parameters –  $\{\beta_0, \beta_1, \beta_2\}$  in the case of  $f_1$ , and  $\{\beta_k: 0 \leq k \leq 2\}$  in the case of  $f_2$ .

29} in the case of  $f_2$  – fixed by the first data set. In this case, we get a value smaller than  $6 \times 10^{-31}$ , indicating immensely more discrepancy between the new data and what  $f_1$  predicts than between these data and what  $f_2$  predicts.<sup>3</sup> Thus, the extra structure in the curve given by  $f_2$  errs in that it captures much variance in the data that is unrelated to the true relation between X and Y.<sup>4</sup> In short, a bizarre model like  $f_2$  that captures *all* the (original) data is clearly *inferior* to the far more standard model like  $f_1$  that doesn't. (As a bit of terminology, I will use 'model' and 'theory' interchangeably.) In particular, the simpler model does a massively better job at predicting the general trends of new data as it arrives.

What then is the relation between  $f_1$  and the actual raw data? This relation is given by adding a 'residual' or 'error' term to our equation:

$$(7) \quad Y_i = f_1(X_i) + \varepsilon_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \varepsilon_i,$$

The term  $\varepsilon_i$ , whose value varies as  $i$  varies, expresses whatever deviation is present between  $f_1$  and the raw data. As (7) shows,  $\varepsilon_i = Y_i - f_1(X_i)$ . In practice, scientific models of complex phenomena *never* perfectly fit the data, and there is *always* a residual element ( $\varepsilon_i$ ) present. This is so even when the system under study is completely deterministic. E.g., the true model might be something like

---

<sup>3</sup> From a God's-eye view, this is unsurprising, because  $f_1$  is the form that actually generated the data. I used the formula  $Y_i = 3 + 4X_i + 2X_i^2 + \varepsilon_i$ , where  $\varepsilon$  and  $X$  were normally distributed with a mean of zero and variances of 100 and 10 respectively.

<sup>4</sup> There's much more to be said about the general issues of model construction and model selection; cf. e.g. Forster and Sober 1994, Burnham and Anderson 2002 for further relevant discussion.

$$(8) \quad Y_i = f_1(X_i) + f_3(Z_{1i}, \dots, Z_{ki})$$

In such a case,  $Y$  is always an exact function of  $X$  and  $Z_1, \dots, Z_k$ . However, the influence of the  $Z_j$ s may be very small, very complicated, unknown, poorly understood, etc. Thus, for any number of reasons, it may be natural to model the phenomena with  $f_1$ , all the while realizing that the existence of residuals in the raw data show that there is more to the full story than is presented by  $f_1$ . (In fact, residuals may correspond roughly to the philosophical notion of a ‘*ceteris paribus*’ clause.<sup>5</sup>)

There’s nothing more basic to statistical research than the idea that the best (or true) theories/models will imperfectly fit the actual data. Indeed, that’s why statistical research is founded upon probability theory instead of directly on algebra and analysis. In other words, in real empirical research of any complexity, there will always be unsaved phenomena. But this is not a criticism of statistical modeling. Rather, it is a reflex of the fact that actual data is frequently the result of multiple influences, only some of which are relevant for a given project.

#### 1.4 Is Linguistics Relevantly Different from the Other Sciences?

Let’s get back to semantic theorizing. Notice that like  $f_1$  and  $f_2$ , semantic theories are theories of a complex phenomenon (i.e., the interpretation of language). The raw data of a sample of the linguistic phenomena aren’t numerical; instead, they are assessments about certain types of

---

<sup>5</sup> The correspondence may not be perfect, though, since in real life as well as in the mathematical assumptions underlying this part of statistical modeling, the probability that the residual contributes nothing to the equation is zero. Thus, it may not be true that *ceteris paribus*,  $Y_i = f_1(X_i)$ , depending on what one’s theory of *ceteris paribus* clauses is.

idealized<sup>6</sup> linguistic behavior: what sorts of things would typical speakers communicate by uttering a given sentence, and under what conditions? That is, the raw data of semantic theorizing are the intuitive truth conditions of our utterances, as we do or would make them in various contexts. Proceeding like the statistical researcher, the Semantic Minimalist begins by hypothesizing that there is some relatively simple structure – i.e. simple in comparison to the complexity of the raw data – that accounts for much of the collective behavior of the raw data. In order to obtain this simple, general structure, some aspects of the raw data (i.e., some aspects of the intuitive truth conditions) must be ignored, just as we ignore some aspects of variance in the statistical case. In general, both minimalist semantic theories and scientific models have the same general form:<sup>7</sup>

$$(9) \quad \text{Raw Data} = (i) \textit{Effects} \text{ of processes under study } (ii) \textit{Interacting} \text{ in some way with} \\ (iii) \textit{Residual} \text{ Effects}$$

The minimalist theory supplies some aspects of meaning that are hypothesized to capture much of the general behavior of the *totality* of the data set. By assumption, the outputs of this theory are not assumed to capture all of the raw data (i.e., intuitive truth conditions of utterances). In fact, it is not even assumed that the semantic theory will *ever* capture all of the intuitive truth

---

<sup>6</sup> The notion of idealization in linguistics and the other sciences has been discussed at great length in many places (e.g., Liu 2004, Chomsky 1986, and citations therein). Since the primary data of interest here concern ‘intuitive truth conditions’, the idealizations at play here are substantially less (although by no means absent!) than in other areas of linguistics.

<sup>7</sup> Of course, correlation does not imply causation, so more is needed here than just the regression analysis. Similar issues apply to linguistic theories as well. For simplicity’s sake, I will ignore these matters, and assume that both types of models support the scientific interpretation in (9).

conditions. As we've seen, such an outcome is absolutely standard science. Our pollution researcher would not assume that there will be some raw datum  $Y_i$  such that  $Y_i = f_j(X_i)$ , with no contribution from the residual effects. Indeed, it is quite typical to expect that  $\varepsilon_i$  will never equal 0, particularly when the phenomenon under study is extremely complex. (When the phenomena are quite complex, a model may be considered significant even if it captures as little as 16% of the raw data (e.g. R. Putnam 2000, 487).) Likewise, the intuitive truth conditions of utterances may always be determined by both the minimalist theory of meaning, and by other interacting aspects of communication. These other aspects of communication are familiar: background beliefs, indexical-fixing elements, demonstrations, 'performance' capacities of speaker/hearers, etc.

As an aside, notice how Chomskian the present view is. Recall that Chomsky often cautions that linguistic theories are not obliged to capture all the facts about various grammaticality judgments, or all of various details present in collections of data. By seeking out more general patterns, we may be able to learn about a speakers' linguistic 'competence', which can be masked by additional 'performance' factors that are also realized in the empirical data. For instance, Chomsky writes:

Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogeneous speech community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance (Chomsky 1965, 3)

By idealizing away from various extraneous factors, we can ‘smooth’ out the raw data of linguistics and thereby hopefully reveal the more significant forces and structures underlying human linguistic abilities.

To return to our main theme, a minimalist semantic theory is a *theory* about the nature of the raw data. Like any other scientific theory, one of its essential rights and obligations is to characterize those parts of the raw data it considers to be truly part of the phenomenon under study, and what other parts are due to extraneous processes; cf. (i) and (iii) in (9). (Several researchers appear to have missed this point. In addition to those cited above, cf. Allen 2003 (552); Stone and Davies (2002, 285 – 7) provide a lucid discussion of this aspect of Chomsky’s reply to Davidson 1986.) The fact that semantic theories get to characterize their own scope also means that they should be judged by the standard – complicated but familiar – criteria of successful scientific theories: simplicity, elegance, predictive fecundity, integration with other successful theories which collectively account for the raw data (or, more typically, hopefully someday will account for the raw data), etc. Methodologically speaking, demanding that a semantic theory sometimes exactly characterize the intuitive truth conditions of utterances appears to be just like demanding that statistical models should (at least for some interesting range of values) be like the complex  $f_2$ , instead of like the more standard  $f_1$ . Such a demand would be bizarre and deeply incorrect in the statistical case; I submit it is no better motivated in the case of semantic theorizing.

The points just made show that (AP) and (3iv) place an unwarranted constraint on theory construction. In no other study of complex phenomena would one demand that theories perfectly capture the raw data across some interesting range of cases. (3ii) should be rejected because it is

one of the rights of a theory to provide a theoretically useful characterization of the phenomena it addresses. (3ii) simply denies this when the theory is semantic. Thus, PUP is unsound.

Another way to view the problem here is that PUP depends on an equivocal interpretation of ‘separability’. Everyone can agree that the intuitive truth conditions of our utterances are almost always substantially influenced by pragmatic factors. In this sense, it’s probably true that pure semantic content is ‘inseparable’ from pragmatic factors: in actual language use, you rarely if ever find the former alone, without the latter. This interpretation of inseparability makes (3ii) plausible, but it also undermines (AP) and (3iv). After all, it’s no criticism of a theory that it treats the raw data as being a product of multiple sources. If this is what separability is, then (AP) and (3iv) simply beg the question against minimalist theories.

On the other hand, (AP) and (3iv) are plausible if inseparability means that no reasonable total theory of language will treat minimalist and non-minimalist aspects of meaning as effects of (relevantly) distinct processes. In fact, in order for (AP) and (3iv) to be plausible, the relevant notion of inseparability must require that that all aspects of the intuitive truth conditions be explained by the same theoretical mechanisms. (E.g., (AP) implies that intuitive truth conditions cannot be the result of various processes which can be studied independently of one another, where only some of these processes are lumped together as a ‘semantic theory’.) Now (AP) and (3iv) are virtually tautologies, but (3ii) loses its support. Why should the fact that the intuitive truth conditions of our utterances *do* contain both minimalist and non-minimalist aspects of content be sufficient to license the restriction that any theory of semantic content *must* capture all of these aspects? Such a view clearly begs the question against minimalist semantic theories.

I complete this section by briefly considering an argument in favor of (AP). Recanati writes:

Suppose I am right and most sentences, perhaps all, are semantically indeterminate. What follows? That there is no such thing as ‘what the sentence says’ (in the standard sense in which that phrase is generally used). . . . If that is right, then we cannot sever the link between what is said and the speaker’s publicly recognizable intentions. We cannot consider that something has been said, if the speech participants themselves, though they understand the utterance, are not aware that that has been said. This means that we must accept the Availability Principle. (Recanati 2001, 87 – 88)

Recanati’s claim that most or all sentences are ‘semantically indeterminate’ amounts to the assertion that minimalist semantic theories don’t capture the intuitive truth conditions of most or all sentences. It’s unclear, though, why such a claim should be taken to imply that ‘there is no such thing as ‘what the sentence says’ (in the standard sense in which that phrase is generally used)’. I take it that ‘what the sentence says’ here refers to the content that a (minimalist or other standard) semantic theory ascribes to a sentence. If that is correct, then claiming that there ‘is no such thing’ is simply question-begging. Again, it’s part of the job of a theory to carve out the sub-portion of the phenomenon that it directly deals with, leaving the remaining parts for further theorizing (this theme is revisited in §2). Recanati’s claim that there is no such thing as ‘what is said’ in this context is like saying there is no such thing as the true population model  $f_I$  in the statistical case. Hence, Recanati hasn’t made a viable case for (AP).

In sum, PUP fails primarily because it ignores a basic fact about scientific theorizing: each theory gets to determine what part of the phenomena it addresses, and typically this is only a very proper subpart of the total phenomena. The requirement that a theory accommodate all of the intuitive truth conditions often enough in some relevant range of cases is a restriction on semantic theories that has no precedent in any of the developed sciences. Indeed, it is far more typical to assume that a given theory will not account for the data.

In the final analysis, a lot of the present discussion hangs on the current epistemological situation with respect to linguistics. I've stressed that part of a scientific theory involves carving up the data in one way or another. This aspect of theorizing is especially important in linguistics because of how little we know about how language works. If we understood language better, our additional knowledge would likely constrain what structures, processes, etc. could be plausibly employed by a theory. As it is, though, our lack of knowledge about language leaves us with very few such constraints on theories (at least at the level of detail that is currently relevant here, i.e., whether minimal propositions play some psychological role concerning language).

The difficulty with constraining which entities, etc. a theory may reasonably posit is further compounded by the fact that there is tremendously strong evidence that many cognitive processes, structures, etc. are not introspectively accessible by anyone. Several decades of literature on e.g. human vision and judgment and decision-making have shown that many aspects of these cognitive phenomena are not consciously accessible to us. Similarly, people lack conscious awareness of many hypothesized aspects of language and its processing (cf. Townsend and Bever 2001 for an overview). E.g., people aren't aware of such items as traces, PRO, movement, Merge, etc. that are posited by *some* linguistic theories simply because of the theoretical work they do. But opposing theorists cannot simply announce that the former theories are false, because they employ some entity which they intuitively *feel* is not truly part of language or its processing. The entity in question may well be part of the end product or part of the processing, or both, even if people are not generally consciously aware of it. (Indeed, even in the case of semantics, there is evidence that people have substantial difficulties becoming aware of certain aspects of the meanings of relatively ordinary sentences (Johnson 2006))

In the present case, then, it's not enough for Recanati to simply declare that CL's minimal propositions won't find any place in a scientific theory of language. He also needs to supply enough well-established empirical evidence to show that no such minimal propositions will play any role in a completed correct theory. Lacking such details, CL's theory is not objectionable on these grounds. I take it that this is just a general point about the nature of scientific theories in general. Much of the philosophy of science concerns how scientific theories will often posit unobserved entities, where the justification for treating these entities as real comes only later, as the theory is confirmed by the usual holistic criteria. These are the criteria by which a minimalist theory should be judged, not by a priori speculation about what entities will appear in a completed theory of language.

Finally, it's worth considering the epistemological status of the 'unexplained' residuals of a linguistic theory. (I thank an anonymous reviewer for prompting this paragraph.) One might feel that the residual aspects of a linguistic theory constitute at least *prima facie* evidence against it. That is, if a theory leaves sizeable chunks of the raw data unexplained, then in the absence of any suggestion of how to develop the theory to handle the remaining (aspects of the) data, the theory is rendered thereby less plausible. Of course, we've seen that this is clearly not Recanati's view, which is the topic of this section. Recanati takes the existence of (what I've called) residuals to be decisively disconfirming evidence of a semantic theory. But from an abstract philosophical perspective, Recanati's view seems plausible: it's a problem when your theory won't 'play nice' with other theories to jointly explain the raw data. However, as this paper emphasizes throughout, it's critical to distinguish the evaluative standards applied to those imagined, not-currently-available total theories of a given phenomenon (e.g. language and linguistic behavior), and those standards applied to the highly incomplete and imperfect theories

that we actually have to work with. As Chomsky has frequently noted (e.g. 1992, 214), demanding that a linguistic theory account for all of the raw ‘linguistic’ data would be tantamount to demanding a ‘theory of everything’, which is simply unreasonable at this stage of our understanding of language. Another way to see this last point is that *every* currently available linguistic theory is radically incomplete with respect to the available data. But if no theory comes even close to explaining all the data, then it’s hardly a criticism of one theory in particular that it fails to do so. Unless there is something special about the *way* that a particular theory fails to explain the data (e.g., that it has no hope of ever being developed so as to explain all the data), it’s unclear that the existence of residuals constitutes any serious worry for a theory at all. Rather, the existence of residuals in linguistic theories is, just as in other theories, a mark that our science is not yet complete.

## **2 Dualism II: Defining Theoretical Terms**

I now turn to a second form of methodological dualism. This one concerns the need to provide definitions of the technical or theoretical terms that one uses in linguistic theories. Jerry Fodor (inter alia) has argued that linguists’ failure to do this undermines their ability to appeal to such notions in their theories. I quickly sketch an example of the offending bit of linguistic theorizing, and then we’ll look more carefully at Fodor’s position. We then consider whether Fodor is right that this requirement on theoretical terms holds generally throughout the sciences. It doesn’t. Examining why the requirement doesn’t hold in the (other) sciences explains why it doesn’t hold in linguistic theorizing either.

## 2.1 The Hypothesized Linguistic Structures

A great many linguists hold that there are a small number of ‘thematic roles’ which, for present purposes, we may regard as linguistically primitive semantic elements (e.g., Baker 1988; Dowty 1991; Grimshaw 1990; Hale and Keyser 1986, 1987, 1992, 1993, 1999, 2003; Jackendoff 1983, 1987, 1990, 1997, 2002; Levin and Rappaport Hovav 1995; Parsons 1990; Pesetsky 1995). For example, one such thematic role is that of ‘Agency\*’. The Agent\* of a clause can be thought of as *roughly* something along the lines of the doer of the action described by the verb, if there is such a doer. (Thus, Agency\* is always relativized to a sentence; Bob is the Agent\* of *Bob bought a camera from Sue*, but Sue is the Agent\* of *Sue sold a camera to Bob*, even though these two sentences may describe the same event.) In general, thematic roles are similar but not identical to certain ordinary notions, hence the asterisks.<sup>8</sup>

Agency\* figures into a wide variety of linguistic generalizations and explanations (cf. the citations above). As a simple example, notice that the doer of an action is always in the subject position of a transitive verb (assuming that the verb requires either the subject or object to perform the action):

- (10) a. John kicked the horse.  
 b. Susan kissed the bartender.  
 c. Christine built the shelf.

---

<sup>8</sup> E.g., *the poison* is plausibly the Agent\* in a sentence like *The poison killed the Pope*, even though poison is not an agent of a killing.

Linguists typically hold that this generalization is not accidental, and in fact holds robustly across all human languages. Thus, (11) is normally taken as an important structural generalization about human languages that linguistic theories should respect and explain:

(11) The only verbal position for Agents\* is the subject position.

There is much more to be said about (various theories of) Agency\* and other thematic elements. However, this brief introduction will be enough to introduce the methodological criticism I want to explore.

## 2.2 Criticism of the Theory Via Methodological Principle

Several philosophers have criticized linguists' use of notions like Agency\*. Since Agency\* isn't agency, they have argued that linguists need to define what is meant by this technical term. To give this worry a name, I will call it the *Problem of Undefined Terms* (PUT). The most vocal advocate of PUT is Jerry Fodor. For instance, Fodor writes,

If a physicist explains some phenomenon by saying 'blah, blah, blah, because it was a proton...' *being a word that means proton* is not a property his explanation appeals to (though, of course, *being a proton* is). That, basically, is *why* it is not part of the physicist's responsibility to provide a linguistic theory (e.g. a semantics) for 'proton'. But the intentional sciences are different. When a psychologist says 'blah, blah, blah, because the child represents the snail as an Agent\*...', the property of *being an Agent\*-representation* (viz. *being a symbol that means Agent\**) is appealed to in the explanation, and the psychologist owes an account of what property that is. The physicist is responsible for *being a proton* but not for *being a proton-concept*; the psychologist is responsible for *being an Agent\*-concept* but not for *being an Agent\*-concept-ascription*. Both the physicist and the psychologist is required to theorize about the properties he ascribes, and neither is required to theorize about the properties of the language he uses to ascribe them. The

difference is that the psychologist is working one level up. (Fodor 1998, 59, underlining added; cf. Fodor and Lepore 2005, 353 – 4 for similar sentiments)

Prima facie, Fodor's argument here seems pretty compelling.<sup>9</sup> If *Agency\** doesn't mean agency, then what does it mean? If you don't know what *Agency\** means, then how can you use it in an alleged 'generalization' like (11)? If you only say that the subject of transitive verbs can be the *Agent\** of the verb, but no independent constraints are placed on what it is to be an *Agent\**, then the alleged generalization about verbs has no content. To see this, just replace the technical term *Agent\** with any other made-up word, say *flurg* (cf. Fodor 1998, 59). Now (11) can be restated as *The only verbal position for flurfs is the subject position*. Obviously, with no theory of what flurfs are, this statement is empty. (The urge in philosophy to demand such definitions appears quite strong. E.g., in an otherwise remarkably subtle and astute methodological discussion, Pietroski (2005 pp. 185ff., 196 (fn. 8), 202) similarly criticizes many linguists' use of the notion of 'direct causation'.)

### 2.3 Is the Principle Justified by General Scientific Considerations?

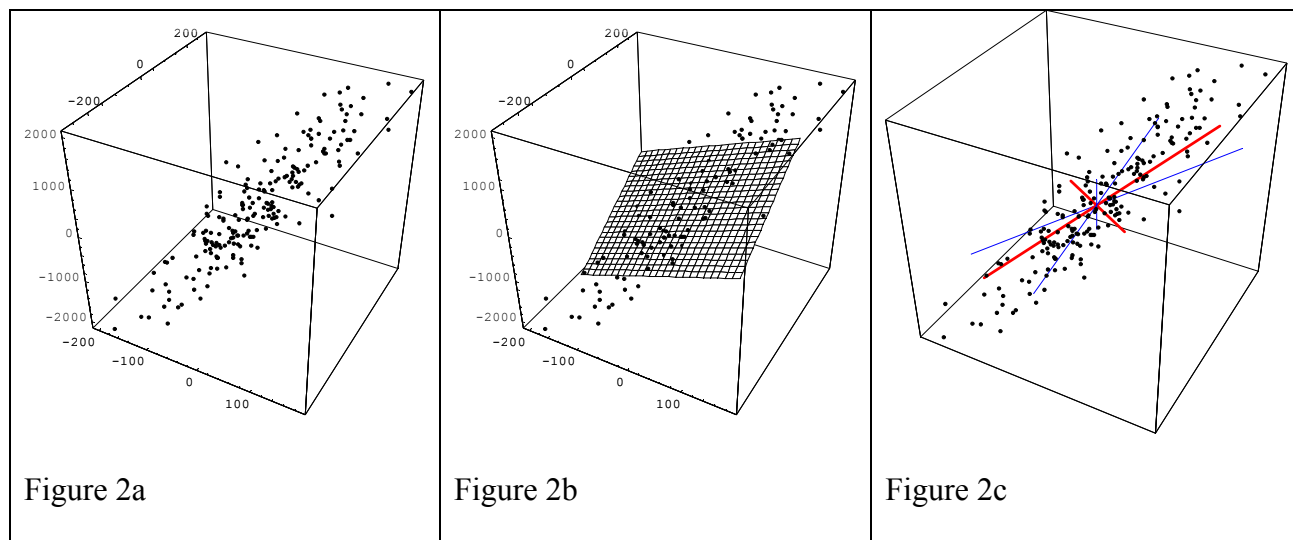
As his allusion to physics makes clear, Fodor believes that any scientific theory must 'provide an account of what property' is denoted by any theoretical term it uses. Indeed, his criticism is just that linguistics violates this general principle of science. Alas, this principle is not generally true

---

<sup>9</sup> It is somewhat odd that Fodor endorses PUT. After all, the linguist posits the notion of an *Agent\** as a linguistically primitive element, and Fodor has vigorously defended the view that (the concepts denoted by) such linguistically primitive elements can't be defined. Instead, he maintains that the best theory of what they mean is simply given atomistically. According to atomism, the best theory of meaning says only that the word *dog* means dog (or better, *dog* denotes the concept of dogs). One would think that such an attitude would carry over to other parts of language, too: *Agent\** means *Agent\**.

in the sciences. Theoretical terms are frequently introduced to denote some hypothetical property that is posited in the theory in order to account for some kind of ‘surprise’ or ‘pattern’ in the empirical data. Moreover, the nature of this theoretical property is often determined not by stipulation in advance, but by continuing theoretical work. This practice is especially common in the earlier, developing stages of some area of inquiry, which is undeniably where all areas of linguistics currently are at. In short, it is common for a scientist to ‘theorize about the properties he ascribes’, but this does not amount to producing more of ‘an account of what property’ he postulates than the linguist produces regarding Agency\* and the like.

To flesh out these ideas, let’s consider an example. As before, I stress that this is only an example. It is meant to illustrate what scientists standardly do when exploring complex phenomena. Suppose we are examining the concentrations of three chemicals X, Y, and Z in a given region. One hundred groundwater samples are taken from the region, and the amounts of each of X, Y, and Z are recorded. When the data are plotted as points on three axes, they are distributed as in Figure 2a below. If the concentrations of the three chemicals were completely uncorrelated, we would expect the data to form a sphere inside the cube. But rather than being randomly dispersed in this sense, the data collectively display a general ‘pattern’. This structure is of course a real surprise, since it’s extraordinarily improbable that a random sample of unrelated measurements would ever yield such a pattern. (The boxes are scaled to a 1-1-1 ratio to visually present the correlations, as opposed to the covariances, of the three variables.)



It's the essence of the sciences not to ignore such patterns in the world. A natural first step is to try to understand 'how much' of a pattern is there, and what its nature is. Obviously, the relative concentrations of  $X$ ,  $Y$ , and  $Z$  are related. From the geometric perspective of the cube, the data appear to be organized around an angled plane, depicted in Figure 2b. The fit isn't perfect, and the planar surface lies at a skewed angle, so all three axes of the cube are involved. But if we used a *different* set of axes, we could view the data as organized primarily along two axes. In other words, suppose we replaced axes  $X$ ,  $Y$ , and  $Z$  with three new axes,  $A$ ,  $B$ , and  $C$ . (If we keep  $A$ ,  $B$ , and  $C$  perpendicular to one another, we can think of ourselves as holding the data fixed in space, but rotating the cube.) Moreover, suppose we choose these axes so that the  $A$  axis runs right through the center of the swarm of data. In other words, let  $A$  be that single axis on which we find as much of the variation in the data as possible. If we had to represent all the variation in our data with just one axis,  $A$  would be our best choice. It wouldn't perfectly reproduce all the information about  $X$ ,  $Y$ , and  $Z$ , but it would capture a lot of it. Suppose that we now fix the second axis  $B$  so that it captures as much of the remaining variation in the data as possible, that is unrepresented by the  $A$  axis. Together,  $A$  and  $B$  determine a plane lurking in the

three-dimensional space. (The two darker lines in Figure 2c correspond to Axes A and B.) By projecting all the data onto this plane, we could recover most, but not all, of the information about the variation in the data. (We miss exactly that information regarding how far to one side or another from the plane the actual data points lie.) If we set axis C to best capture the remaining information, we then recover all the information in the original space. It is more common, though, to represent the data using fewer axes than the data were originally distributed on. If, say, we are satisfied with the amount of information we get using only one or two axes, we can represent the data in a less complex manner, using only one or two dimensions, instead of the original three-dimensional format. Whether we represent the pattern in the data with just axis A, or with axes A and B is not a decision that the mathematical analysis itself makes for us, although in many cases other techniques or considerations will provide strong evidence for one option.

The technique just described is called Principal Component Analysis (PCA). The mathematical essence of PCA involves finding new axes on which to represent the data. The first Principal Component (PC) is axis A, and the second and third PCs are B and C respectively. Clearly, this technique is not restricted to three dimensions – it can be used with any (finite) collection  $X_1, \dots, X_n$  of measurements.<sup>10</sup> The real scientific import of PCA comes when we find

---

<sup>10</sup> Briefly, here is how PCA works. Start with an  $n \times n$  correlation (or covariance) matrix, where the  $ij$ th entry gives the correlation between  $X_i$  and  $X_j$ . Then extract all the eigenvalues (there are typically  $n$  of them) and eigenvectors of unit length from this matrix. Ordering these eigenvectors according to the size of their eigenvalues, we obtain our PCs (cf. Basilevsky 1994 for proofs). The  $k$ th eigenvector gives the direction of the  $k$ th axis in the  $n$ -dimensional data space. Moreover, the  $k$ th eigenvalue expresses the amount of total variation in the data captured by the  $k$ th PC. (When working with a correlation matrix, the total variation will always be  $n$ .) Also, the amount of variation in a given measurement  $X_i$  that the  $k$ th PC accounts for is given as  $\sqrt{\lambda_k} a_{ik}$ , where  $\lambda_k$  is the  $k$ th eigenvalue and  $a_{ik}$  is the  $i$ th element of the  $k$ th eigenvector. This form of PCA produces perpendicular axes, each of which accounts for the

that e.g., one or two PCs can account for say 95% of the variation in one or two hundred types of observations. Geometrically, this is like finding that in a space of one or two hundred dimensions, all the data are arranged almost perfectly in a straight line or on a single two-dimensional plane lurking in that space. Such a pattern is far too extreme to be random, and it cries out for explanation. PCA and related techniques can help quantify this pattern in useful ways.

By exposing a small number of dimensions of variation which capture most of the variation in the data, we derive an *explanandum*. Why should the concentrations of three – or more realistically, 30 or 100 – different chemicals behave as though they were from only two sources? At this point, an empirical/metaphysical hypothesis suggests itself: maybe they behave this way because there are two sources responsible for the emission of the chemicals. At this point in the investigation, we may not be able to say much about these two hypothesized sources other than that they are what are emitting the chemicals in question. We can't, for instance, automatically infer that one source corresponds to axis A and the other source to B. Axes A and B determine a plane in the data-space, but infinitely many other pairs of lines (not necessarily at right angles) could also determine that same plane; cf. Figure 2c. If the two sources do correspond to axes other than A and B, then they each emit different concentrations of X, Y, and Z than A and B predict. If the sources correspond to two axes that are not at right angles, this means that their emissions of chemicals are correlated with one another. Finally, even the notion of a 'source' must be understood in a broad functional sense. There may be two physical sources that both emit the same relative concentrations of X, Y, and Z, and so one of the axes represents

---

maximum amount of remaining variance in the data. Once one decides to retain  $m$  PCs (where  $m$  is almost always much less than  $n$ ), the basis for the  $m$ -dimensional subspace can be changed to suit background hypotheses, etc.

both their emissions. Alternatively, one axis could represent the joint effects of several physical sources that emit different concentrations of X, Y, and Z, but which are all perfectly correlated with one another. At the same time, the hypothesis that there are two sources of emission is a very strong and testable hypothesis, not least because together they must nearly determine the particular plane in the data-space (cf. e.g., Malinowski 2002, ch. 10 – 12 for discussion of this and other uses of PCA and related techniques in chemistry).<sup>11</sup>

PCA is one of a large family of methods – which includes factor analysis, structural equation modeling, latent class analysis, discriminant analysis, multidimensional scaling, and others – for exploring the extent to which there is latent structure in the data. These techniques all involve uncovering underlying regularities that appear when individuals (persons, groundwater samples, etc.) are measured in a variety of different ways. (Such techniques bear some similarity to Whewell’s (1840) ‘consilience of inductions’, although they supply much more information and structure.) In the study of complex phenomena, where many different sorts of measurements are possible, the use of such techniques is extremely common, particularly in the early stages of inquiry, but also consistently throughout the development of the theory. Indeed, one would be hard-pressed to identify a line of scientific inquiry into some complex phenomena where such techniques *weren’t* used. It’s just what you do with data.

In short, in the early stages of ordinary scientific inquiry, it is perfectly possible to hypothesize the existence of unobserved empirical structures, sources, etc. without having much of a theory about their natures. Of course, as the investigation develops, the chemist ‘owes an

---

<sup>11</sup> There are *lots* of other ways to explore the results of the PCA. If two candidate sources are found, the PCA will supply evidence about their correlation. If the sources are different factories, this may indicate e.g., the degree to which they are working together, or are both influenced by the same economic factors, etc. Also, if the two sources are discovered using only some of the measurements, say X and Y, then the PCA will express the relative amounts of Z that the sources emit. If Z is a noxious pollutant, this may be extremely important information.

account' of the nature of the sources of emission. However, at the early stages of inquiry, the details of this account may be a long ways off. Indeed, for highly complex situations, there may be a lengthy initial period of many rounds of data analysis, with a focus only on figuring out how many latent structures there are and what kinds of overt measurements they affect. If scientists were required to precisely characterize all hypothesized structures, a great deal of successful research into complex systems would be illegitimate.

In the next section, we will see that positing unobserved linguistic structures in linguistics works by recapitulating point-by-point the underlying logic of uncovering a few dimensions of variation from a larger-dimensional data-space. Just as with PCA, we'll see that it is the 'multivariate' nature of linguistic evidence that plays a crucial supporting role in the hypothesis of unobserved linguistic structure.

## 2.4 Is Linguistics Relevantly Different from the Other Sciences?

If we consider the underlying logic of quantitative methods like PCA, here is what we find. Sometimes there are just a few significant PCs present in a high-dimensional data-set. The fact that the data are largely clustered in these few dimensions is a significant explanandum that needs explaining. This need can justify the provisional adoption of a *hypothesis* that there is some unobserved empirical structure underpinning the PCs. Of course, the hypothesis may be wrong, and even if it is on the right track, much of the nature of these unobserved structures remains to be discovered. Crucially, though, the multivariate nature of the PCs constrains (and thus helps to form) hypotheses about the unobserved structures responsible for the PCs. That is, the fact that the PCs are built out of *multiple* overt measurements severely limits what sorts of things they could represent. E.g. inspection of our simple example above shows that if you know

what a proposed PC predicts on just one dimension, say X, then you can tightly constrain what it predicts on the remaining dimensions. Put another way, not all sets of possible predictions correspond to possible PCs.

The situation with linguistics is very similar to what we have just seen with quantitative data analysis. For starters, the discipline of linguistics is certainly still at an early ‘exploratory’ stage, and it certainly concerns a very complex phenomenon. Moreover, aspects of these phenomena are represented with various *measurements*. Measurements in linguistics typically are not quantitative, but instead concern such things as the grammaticality or acceptability of a sentence, its sound and meaning, etc. These measurements are used to reveal and explore various patterns that exist within various interestingly clustered sets of sentences. The goal is to uncover the latent structures responsible for (much of) these patterns. Moreover, just as with PCA, the structures we hypothesize may not capture all the empirical data, since linguistic theories will involve residual effects, as we saw in §1. Thus, it’s natural to understand thematic roles like Agency\* as having the same sort of epistemic status as the latent structures in any other exploratory data analysis. True, we don’t know fully what Agency\* is, but that doesn’t mean that we can’t provisionally hypothesize the existence of such a latent element as part of a theory about why our overt measurements (i.e., linguistic judgments) behave as they do. In the early stages of inquiry, one chooses to provisionally hypothesize the existence of thematic roles or of a correlate of a highly significant PC for largely the same reasons: both types of hypotheses are testable in many ways and have lots of room for potentially wide-ranging augmentation and refinement through further scientific inquiry.

So why did the Problem of Undefined Terms seem so compelling? I suspect that there are two main reasons for this.<sup>12</sup> First, PUT encourages us to think of a *completed* theory of Agency\*. In a finished linguistics, we would expect more details about Agency\* than are currently available (for some recent empirical exploration of direct causation, cf. Wolff 2003, Johnson 2006). But the present issue concerns justifying the *very beginnings* of such a theory. Thus the relevant immediate question is not whether Agency\* is a real linguistic element. Instead, it is whether the notion of Agency\* is sufficiently promising as a part of a linguistic theory (given the other available theories and whatever else we know) that there is merit in provisionally hypothesizing its existence and exploring the consequences for the resulting linguistic theories.

Second, recall that the real worry behind PUT was that a technical notion like Agency\* is too unconstrained and underdetermined to be useful in theorizing or in forming generalizations. If there are no constraints on what it is to be an Agent\*, then one can make a generalization like *All (or only) Agents\* are Fs* true by brute force. For any potential counterexample C to the generalization, nothing prevents you from stipulating that (e.g.) C is not an Agent\*. But this worry is defused when we notice that, just as PCs can only be (non-trivially) extracted from collections of more than one sort of measurement (e.g. X, Y, and Z concentrations), the extraction of linguistic structure always involves multiple sorts of linguistic phenomena. This

---

<sup>12</sup> Actually, there are probably some further reasons, which have to do with the unclarity of some of the crucial judgments (cf. Schütze 1996 for sustained discussion), and the fact that such unclarity often seems to accumulate as theories become increasingly complex. However, since these issues do not pertain directly to the Problem of Undefined terms, I will leave them for another day.

last point, and several others, can be elucidated by considering a simplified sketch of how one might justify positing Agency\* in a linguistic theory.<sup>13</sup>

Suppose a linguist, call her Lana, notices that some derived nominals (i.e., nouns that Lana hypothesizes to be derived from an underlying verb) have a form that corresponds to the passive form of the verb (12), but other verbs do not (13):

- (12) a. Sharon proved the theorem/ the theorem's proof by Sharon  
 b. John destroyed the vase/ the vase's destruction by John  
 c. John created the vase/ the vase's creation by John
- (13) a. Sue loved Mary/ \*Mary's love by Sue  
 b. Sue resembled Mary/ \*Mary's resemblance by Sue  
 c. Sue awakened / \*the awakening by Sue

After studying this pattern, Lana begins to explore the hypothesis that there is some structural property she calls *flurg* present in the nouns or verbs in (12) and absent in (13), or vice versa. At this point, *flurg* simply encodes a difference between two kinds of words. But then Lana notices that with nominals like (12), the nominal can be the complement of a possessive, or it can have a passive *by*-phrase adjunct, but not both (although it can take some *by*-phrases and possessives):

- (14) a. The Romans' destruction of the city

---

<sup>13</sup> Since I only want to illustrate a very common method in linguistics, I omit lots of details and data, in order to avoid the complexities of doing linguistics straight out. You needn't be convinced of the example's details in order to understand the method employed. All the ideas and data, though, are very familiar from the linguistics literature.

- b. The destruction of the city by the barbarians
- c. \*The Romans' destruction of the city by the barbarians
- d. The Romans' destruction of the city by catapults and mass attack

Lana now hypothesizes that such nominals have some property that can license either the possessive or the passive *by*-phrase, but not both. At this point, Lana makes a crucial conjecture: she hypothesizes that this property is *flurg*, the very same one used in (12) – (13). As a third bit of data, Lana notices that languages lack symmetric pairs of verbs for asymmetric events. E.g., while we have verbs like *kick*, *lift*, *build*, etc., we don't have verbs like *blik*, where *x blik*s *y* if and only if *y* *kick*s *x* (and similarly for *lift*, *build*, etc.).<sup>14</sup> Once again, Lana attempts to reduce the complexity of her theory: she hypothesizes that *flurg* is once again responsible for this phenomenon. Thus, she puts forward that *flurg* must necessarily be located in the subject position of the verb, and can never appear in object position. Of course, any of these hypotheses could turn out to be false, but so far Lana feels that her developing theory of *flurg* and its roles in language is sufficiently plausible to merit further study.

As Lana continues to study the words that she hypothesizes have *flurg* versus those that don't, she begins to sense that words with *flurg* all share a certain semantic similarity, although she cannot fully articulate what it is. To explore this hunch, she gives a very brief characterization of what little she knows about this semantic similarity to a variety of people (both linguistically trained and untrained). She uses only a couple of words as examples, and then gives her subjects a large number of other words, and asks them to indicate whether they

---

<sup>14</sup> In contrast, notice that we do find symmetric pairs elsewhere; e.g., the direct and indirect objects of *Sue sent a letter to Tim* and *Sue sent Tim a letter* can be exchanged without any apparent change in meaning.

perceive this hypothetical semantic property in the expressions, and if so, where (e.g., in subject or object position). She discovers an enormous amount of agreement across her subjects. They typically find flurg clearly present or absent in the same places, and are unsure about the same cases. As Grice and Strawson (1956) noted, this kind of agreement marks a distinction (between the presence and absence of flurg), even if many details about the nature of the distinction are unknown. Attending to the cases where there are uniformly clear judgments, Lana notices that these are also places where the other hypothesized effects of flurg occur. Thus, she further expands her hypothesis about flurg, claiming that it is, or is associated with some kind of semantic feature, which she is currently investigating, but has not yet fully identified or even confirmed. Since this hypothesized semantic feature of flurg corresponds roughly (albeit not completely) to the notion of agency, she renames *flurg Agent\** out of convenience.

Notice that in the story just told, the nature of flurg is unconstrained only at the very beginning, when it is hypothesized to underwrite just one distinction. However, as the theory is developed to account for multiple distinctions, flurg becomes more tightly constrained. By the end, the hypothesis that flurg exists is rather demanding. It is not enough for subjects to simply feel that flurg is clearly present/absent in a given word; the theory also makes (heavily *ceteris paribus*, as it turns out) predictions about the syntactic behavior of that word, and it predicts where in the word, if anywhere, subjects will sense flurg's semantic presence. The multiple dimensions that are used to characterize flurg render it anything but an unconstrained hypothetical element.

In PCA, one seeks out a few vectors in the data space that explain most of the variation in the data. In linguistics one seeks out a few structural elements that explain most of the variation in the data. In neither case are these elements always fully defined or understood. Demanding a

complete definition in the linguistic case is a methodologically dualistic standard. I know of no defense of such an atypical stance towards linguistics. Certainly Fodor's incorrect appeal to the sciences (quoted above) offers no such support.

As a final comment, it's worth observing that the parallel in linguistics with techniques like PCA is quite strong. Indeed, work in consensus theory suggests that it might be possible to perform some form of latent variable analysis on the kinds of collective linguistic judgments discussed above (e.g., Batchelder and Romney 1988). In such a case, the evidence for a structural element like Agency\* could be reduced to something like the question of whether the first PC in the data space captured nearly all the variation in the data. This topic is explored elsewhere.<sup>15</sup>

### 3 Dualism III: Aggregation and Degrees of Accuracy

So far, we've examined two dualistic principles that make linguistic theorizing *harder* in some respects than ordinary scientific theorizing. Could there also be some commonly accepted aspects of linguistic methodology that make linguistic theorizing *easier* than scientific theorizing? In fact, there are, and in this section, I gesture at some of them.

Before beginning, two caveats are in order. First, I illustrate the divergences between linguistic and (other) scientific methodologies with a particular example. But those familiar with

---

<sup>15</sup> It's also worth noting that, just as with quantitative methods, one can also employ linguistic theorizing from the standpoint of mere 'curve fitting'. Sometimes unobserved numerical structures like PCs are extracted and used without any effort to find an empirical interpretation of them. Instead, their value comes solely from their practical utility in reducing the complexity of the data, allowing it to be more easily understood and manipulated. Although a discussion of this issue would take us too far afield, it is worth noting that in linguistics too, hypothesized structures may serve the legitimate and useful purpose of allowing linguistic data to be more easily managed, without worrying about their psychological interpretation.

mainstream linguistic methods will recognize that the morals of this case study generalize very broadly to a vast amount of linguistic research. Thus, there is nothing special about the particular example I use. Also, since the purpose of the example is only to illustrate certain widely used *methods* of linguistic theorizing, I will not try to exhaustively characterize the relevant literature. For present purposes, that would only obfuscate matters. (Indeed, the precise details are unimportant enough that readers familiar with linguistic methods may wish to skim the example, and go right to §3.2.)

Second, my critical remarks are not intended to be some sort of ‘tearing down’ of linguistic theory or the ‘harassing of emerging disciplines’ (Chomsky 2000, 60, 77). Rather, I suggest only that linguistics should follow the other sciences, which routinely study their own methodologies in order to better understand, use, and improve them.<sup>16</sup> (There have been some honorable efforts in linguistics to do just this, e.g., Schütze 1996, Christie et al. 1999, Prince 2002a,b. However these authors do not address the issues discussed below.) Without such study, there will continue to be many reasons why Chomsky was incorrect to claim that ordinary linguistic practices have ‘exhausted the methods of science’ (1986, 252).

### 3.1 An example of linguistic theorizing

In a series of papers, Norbert Hornstein (1998, 1999, 2000, 2003) argues that the phenomenon of linguistic ‘control’ can be accounted for simply by allowing movement into theta positions. E.g., the relevant syntax of (15a) does not have the traditional form in (15b), where PRO is a distinct lexical item controlled by *Sue*. Instead, the proper form is in (15c), where *Sue*

---

<sup>16</sup> To verify this last claim, one need merely consult a current statistics journal, or a journal of a particular science that publishes papers on mathematical methods (e.g., *The Review of Economics and Statistics*, *The Journal of Mathematical Psychology*, *Psychometrika*, *Econometrica*, *Biometrika*, etc.).

has moved from the lower subject position to the higher one. (Following Hornstein, I treat movement as a combination of the Minimalist operations of Copy and Merge.)

- (15) a. Sue wants to win;  
 b. Sue<sub>i</sub> wants [PRO<sub>i</sub> to win];  
 c. Sue wants ~~Sue~~ to win.

More generally, Hornstein holds that linguistic theories do not require the null pronominal element PRO or its associated ‘control module’ that determines the referent of an occurrence of PRO. We don’t need these things, Hornstein argues, because the phenomena that initially motivate positing them can be accounted for by appealing to independently motivated components of the grammar. Movement (aka Copy and Merge), Hornstein assumes, is a prevalent feature of grammar. If all the relevant facts can be accounted for without positing PRO, then *ceteris paribus*, linguistic theories should favor the simpler theory and reject the employment of PRO. In the development of this theory, Hornstein also notes several advantages. Here are two representative examples.

It’s well-known that both PRO and traces (i.e., the residue of Copy and Merge) are phonetically null. By identifying PRO and traces, we reduce the number of ‘unobserved’ structures in need of explanation/defense from two just one. Similarly, the fact that *wanna* contraction can occur either with raising or with control will now require only one explanation:

- (16) a. I seem to be getting taller.  
 a’. I seemta be getting taller.  
 b. I want to get taller.  
 b’. I wanna get taller.

Famously, there is some need to explain this type of contraction, because it does not happen willy-nilly. You can turn *You want to help Mary* into a *wanna*-question by asking *Who do you wanna help?*, but you can turn *You want John to help Mary* into a question only by asking *Who do you want to help Mary?* It is not grammatical to ask *\*Who do you wanna help Mary?*

As a second advantage, Hornstein considers ‘hygienic’ verbs, as in *Peter washed/dressed/shaved Dan*. These verbs are interesting, because they can also appear intransitively:

(17) Peter washed/dressed/shaved

Notice that (17) has a reflexive meaning; it says that Peter washed (dressed, etc.) *himself*. This contrasts with other verbs that can drop their objects; *John ate* does not mean John ate himself, only that he ate something. According to standard views, the reflexive behavior of (17) is puzzling, since a reflexive reading would most naturally be supplied by an element like PRO, but PRO is typically thought to appear only as the subject of a clause (e.g., Chomsky and Lasnik 1993).<sup>17</sup> But on Hornstein’s view, the syntactic element PRO is replaced by whatever structure underlies movement phenomena, and it is well-known that movement can occur from object position – e.g., on Hornstein’s view, *who did Shaun kiss ~~who~~<sub>i</sub>* would be an example.<sup>18</sup> Hornstein

---

<sup>17</sup> For instance, *Bill wants Mary to kiss* cannot have the structure *Bill<sub>i</sub> wants Mary<sub>j</sub> to kiss PRO<sub>ij</sub>*; it can mean neither that Bill wants Mary to kiss Bill nor that Bill wants Mary to kiss herself.

<sup>18</sup> On a syntactic note, it is standard to distinguish (roughly speaking) A-movement from A’-movement. However, if one relinquishes, as Hornstein does, the constraint that a syntactically realized NP (or DP, I won’t adjudicate here) must bear at most one theta role, such a distinction becomes less well motivated. As Hornstein discusses at length, a central component of his theory is that syntactic chains can bear multiple theta roles, a position earlier explored by

shows how the reflexive readings in (17) are (relatively) neatly and unproblematically produced within his theory.

Unsurprisingly, Hornstein's proposal has not gone unnoticed (e.g., Brody 1999, Culicover and Jackendoff 2001, Landau 2000, 2003, Manzini and Roussou 2000). Here are two representative criticisms of the view. The first problem comes from Landau (2003), who argues that Hornstein's theory has problems accounting for 'partial control', illustrated in (18):

(18) The chair of the department wanted to meet on Tuesday afternoon.

(18) is most naturally interpreted as expressing that the chair wanted some set *X* to meet on Tuesday afternoon, where *X* contains the chair and at least one other person. Roughly and intuitively speaking, partial control constructions are distinctive in that the controlling DP is only a proper subset of the collective subject of the embedded clause. That is, (18) does not mean that the chair wants herself (and herself only) to meet (herself) on Tuesday. Further evidence that there is a plural syntactic subject in the lower clause comes from the ability of partial control to support distinctively plural types of predicates and anaphors:

- (19) a. Susan enjoyed getting together on weekends.  
 b. Steve wondered whether helping one another would be productive in the long run.

---

Williams (1980, 1983). This relaxation of the Theta Criterion in many ways is a central bit of machinery of his theory.

It is hard to see how a ‘control as movement’ view such as Hornstein’s can handle partial control. According to him, the relevant structure of (18) is simply *The chair<sub>i</sub> wanted to ~~the chair<sub>i</sub>~~ meet on Tuesday afternoon*. Nothing in Hornstein’s theory appears to explain how the overt copy of *the chair* denotes a single person, but the deleted copy of that very expression denotes a group, of which the chair is only one member. This suggests that the subject of the lower clause of (18) is realized as something other than merely a copy of *the chair*. Further evidence that this other element may be PRO comes from the fact that partial control does not appear to exist in raising constructions, where the syntactic subject of *meet* is typically thought to be a trace (i.e., a copy of the higher subject):

(20) \*The chair seemed to meet on Tuesday afternoon.

‘Without further detail’, Landau argues, ‘one can already see how damaging the very existence of partial control is to the thesis ‘control is raising’. Simply put: *there is no partial raising*. It is not even clear how to formulate a rule of NP-movement that would yield a chain with nonidentical copies’ (Landau 2003, 493).

The second problem comes from Brody (1999, 218 – 19). Consider the following pattern.

- (21) a. John attempted to leave.  
 b. \*John was attempted to leave.  
 c. \*John believed to have left.  
 d. John was believed to have left.

Why can't (21c) be used to express that John believed that he himself had left, just as (21a) expresses (roughly) that John attempted to make himself leave? Similarly, why can't (21b) be used to express that someone attempted to make John leave, just like (21d) expresses that someone believed that John had left? If control is just movement, as Hornstein proposes, then we have no explanation for the different syntactic abilities of what is typically thought to be NP-movement – (21c,d) – and what is typically thought to be control – (21a,b).

Now, of course there is a great deal more to be said about Hornstein's theory. The theory has more prospects and problems than what I've presented, and there are objections and replies to them, and objections and replies to the objections and replies, and so on. But the points to follow can be made by examining just a few considerations.

### **3.2 Aggregating Linguistic Evidence**

In ordinary scientific inquiry, there are many questions regarding the relation between the empirical data, background assumptions, and the resulting theories generated from them. However, current mainstream linguistic practice provides no systematic, principled methods for addressing the vast majority of these questions. Of the many questions that are routinely studied in the other sciences but not linguistics, one is an 800-pound gorilla. The question concerns the aggregation of various bits of one's evidence into a single overall assessment, and the precise estimation of the accuracy of the estimation. I call this issue the *problem of aggregation*. To see what is at stake here, consider the following example.

Imagine a linguist who needs to evaluate Hornstein's view. Perhaps, e.g., she works in a related area of syntax or semantics, and she is trying to decide whether his view is promising enough that she should explore incorporating it into her own theory. (Obviously, if she has little faith in Hornstein's view, she will be disinclined to spend much time and energy on it.)

Simplifying greatly, suppose also that her only considerations about the view concern the advantages and disadvantages just listed: she thinks that (i) Hornstein's theory does a good job accounting for certain reflexive intransitives and (ii) for *wanna*-contraction. But she thinks (iii) the theory is weaker at handling partial control and (iv) passivization. Our linguist recognizes that Hornstein's theory can be made to handle (iii) – (iv), but she also thinks that the only way to do so is rather inelegant and somewhat ad hoc. (For the moment, let's bracket the very difficult issue of how she arrived at these judgments.) Now what does she do? How should she *combine* these individual judgments about Hornstein's theory into a single assessment? At this point, linguistic methodology comes to a grinding halt. If our linguist has no further data or considerations to add, she cannot further analyze the situation except by appealing to her subjective impressions (and perhaps also the impressions of her colleagues) of the overall promise of the theory. In short, the problem of aggregation is that linguistic methodology provides no theoretical tools to guide the inference from a given collection of considerations to an overall assessment of the theory. Similarly, there are no precise, explicit methods to guide assessments of which considerations are more important than which, and of *how much* more important a given bit of evidence is than another bit.

Is there anything wrong with this outcome? Shouldn't the reflective, all-things-considered judgments of professional linguists be trusted about these matters? After all, as several linguists have passionately argued, linguists are highly trained experts whose academic focus is precisely centered on the construction and evaluation of linguistic theories. How could anyone dare question linguists' abilities to aggregate linguistic evidence?

To address this concern, notice that from the fact that linguists know more about the empirical topic at hand than anyone else, it does not follow that the linguist's favored method of

aggregating evidence by making informal, subjective, all-things-considered judgments is the most accurate method. Moreover, issues of the reliability and validity of a decision-making method (of which evidential aggregation is a central case) are empirical ones. In fact, there is substantial empirical evidence that the linguist's method is not the best way to aggregate evidence. The accuracy of large-scale, all-things-considered judgments has been studied intensively for well over fifty years, including the seminal early work of Meehl (1954). The use of these judgments has been rigorously explored in many fields, such as weather forecasting, college/university admissions procedures, psychiatry, clinical medicine, and parole board policies. For instance, Grove et al. examined 136 cases where intuitive all-things-considered judgments were made; only eight of these cases proved to be more accurate than their associated mathematically based prediction rules. This success rate is well within the range for sampling error, suggesting no known successful cases (cf. Grove and Meehl 1996, Bishop and Trout 2002, 2005.) Robyn Dawes summarizes human evidential aggregation as follows:

People are good at picking out the right predictor variables and at coding them in such a way that they have a conditionally monotone relationship with the criterion. People are bad at integrating information from diverse and incomparable sources. (Dawes 1979, 574)

(Indeed, Dawes famously showed that humans are so bad at aggregating evidence that they actually do *worse* than mechanical rules constructed by assigning *random* values for the relative importance of the various types of evidence!)

There is little reason to think that the situation is different in linguistics. Linguists are especially adept at determining what kinds of (linguistic) phenomena constitute evidence for or against a given theory. Moreover, they are amazingly clever at finding such evidence, as well as

new data that present puzzles for every theory. But linguists are also human. And like other humans, scientists included, there is strong reason to believe that their intuitive, all-things-considered assessments of theories are subject to such foibles as overconfidence about the accuracy, success, and promise of their favored theories, and underconfidence about their rivals' theories. As is well known, these normative errors in human judgment and decision-making affect scientists pretty much like non-scientists (e.g., Tversky and Kahneman 1971; Faust 1984; Henrion and Fischhoff 1986; Swets 1996; Swets et al. 2000; Trout 1998, ch. 6; Bishop and Trout 2002, 2005, Nickerson 1998). As Kahneman and Tversky note (1984, 5), the 'stubborn appeal' of these judgmental errors often resemble 'perceptual illusions more than computational errors'.

Leaving linguistics aside for the moment, let's consider how the (other) sciences deal with these weaknesses of human judgment. A glance at the quantitative techniques routinely employed in scientific inquiry shows a strong emphasis on objective methods for the analysis of a model's relations to multiple pieces of the empirical data. As one statistician puts it, the 'most important task [sc. of statistics] is to provide objective quantitative alternatives to personal judgment for interpreting the evidence produced by experiments and observational studies' (Royall, 1997, xi). For example, we earlier (cf. (5) and (6)) considered the sum of squared deviations of the data from the model's predictions:  $\sum_{i \in I} [(Y_i - f_i(X_i))^2]$  In fact, this term yields nothing other than the (squared) distance, in an  $n$ -dimensional Euclidean vector space, between the actual data and the model's predictions. This single quantity represents an aggregation, via addition, of the model's relationship to the totality of a particular data set. Moreover, this aggregation is *perfect* in the important sense that, given the empirical measurements and a candidate model, the Euclidean distance between the empirical facts and the models predictions

is completely determined, with no room for dissent, discussion, or debate.<sup>19</sup> This practice contrasts dramatically with the situation of Lana, whose current methods provide no tools for usefully and uncontroversially analyzing her data.<sup>20</sup>

Importantly, standard evaluative methods account for substantially more features of a model's relationship to the evidence than just the sum of squared residuals. E.g., it is common to modulate the sum of squared residuals by the 'simplicity' or 'dimensionality' of the model, so that the relatively simple model  $f_1$  (with three free parameters  $\beta_0, \beta_1, \beta_2$ ) will be penalized less than the comparatively complex model  $f_2$  (with thirty free parameters  $\beta_0, \dots, \beta_{29}$ ). (The iron workhorse of ordinary workaday empirical research, the analysis of variance, is built on just this idea.) Such techniques are themselves often further developed, resulting in highly sophisticated aggregations of such diverse aspects of a theory as its empirical coverage and its 'simplicity'.<sup>21</sup>

---

<sup>19</sup> Of course, there's always room for debate about *other* aspects of the larger analysis. But as a glance at the actual practice of science immediately shows, even when such techniques provide less than total mathematical guidance, these 'guide rails' nonetheless sufficiently constrain informal speculation so as to supply scientific theories with much of their overall predictive accuracy and reliability.

<sup>20</sup> I've occasionally encountered the assertion that linguists do in fact use the kinds of methods in question. But in the 40 or so works cited herein regarding actual linguistic research into thematic roles and the existence of PRO, exactly zero of them make any use of precise aggregative methods. Similarly, no such methods are used or mentioned in various graduate-level linguistics textbooks (e.g., Haegeman 1994, Culicover 1997). As readers of linguistics journals know, this list could be greatly expanded. Since each one of these works contains many linguistic inferences, we have a sample of literally hundreds of linguistic inferences all involving subjective, informal, all-things-considered judgments. Moreover, the actual use of aggregative results in, say, psycholinguistics or neurolinguistics appears to be quite rare (e.g., Walenski and Ullman 2005).

<sup>21</sup> There are many features of a model besides simplicity that can be important to estimate. For instance, one might want a precise estimate of how 'fragile' the model is to changes in background assumptions, or changes in the data set one used to construct the model. Within the mathematical framework of statistics, such estimations can be had with great precision – and the lack of precision itself can be expressed precisely, as in a confidence interval (e.g. Leamer 1985). Corresponding issues obviously arise for linguistics, and its research would no doubt profit immensely if there were methods for accurately and precisely assessing potential cases of fragility.

For instance, much attention has been paid to various forms of the Akaike Information Criterion (e.g., Forster and Sober 1994, Burnham and Anderson 2002). Additionally, there are a wide variety of other methods of model selection (e.g., Zucchini 2000) but they all centrally involve the mathematical aggregation of various features of the model (e.g., its simplicity, empirical accuracy, coverage, etc.) into a single assessment.

Such additional features of a theory, like its ‘simplicity’, are also important aspects of linguistic theorizing. But how is some aspect of a linguistic theory’s simplicity assessed with respect to the rest of the evidence? How much is the additional simplicity of Hornstein’s theory worth (in contrast, say, to a given theory that posits PRO and a control module)? Is it enough to tip the scales in favor of Hornstein’s theory? If so, how do we know this? Precisely what methods were used to answer these questions about simplicity? How reliable are these methods? How likely are they to accurately assess the theory’s simplicity (in the given respects), as opposed to some other, possibly irrelevant, aspect of the theory? Methods for addressing such questions have been the topic of decades of intensive research in the other fields of empirical inquiry.<sup>22</sup> In mainstream linguistics, however, such questions are rarely if ever addressed with tools other than linguists’ subjective professional judgments.

From a certain perspective, the reliance in linguistics on informal professional judgments is particularly worrisome. Recall the study mentioned above, where the 136 studies showed that informal professional judgment is inferior – often vastly so – to quantitatively based predictions. All of these cases involved judgmental situations where an independently verifiable correct answer was available. But the availability of such information means that most of these

---

<sup>22</sup> E.g., in a one-way analysis of variance, the complexity of the model is given by the number  $c$  of categories the  $n$  individuals are placed in. The variation in the data resolved by the model is multiplied by  $n^{-1}$ , and the remaining variation is multiplied by  $(n-c-1)^{-1}$ .

professionals, during their training and practice up to the time of the study, had access to feedback on whether some of their previous judgments turned out to be correct. Consider e.g. the doctors who aggregated evidence regarding the probability that the patient has breast cancer. They had many opportunities in med school and in their practice to make similar judgments which they later confirmed or disconfirmed. It is well known that this kind of feedback is a powerful aid to improving future judgments. But in the early stages of this difficult field, there are few if any such opportunities currently available to linguists to receive feedback about professional judgments. (E.g., how many theories – at the level of detail as Hornstein’s – have been resolved with the kind of objective certainty as whether a patient develops breast cancer? Of these, how many of them have been used as feedback in training exercises to hone the accuracy of professional linguistic judgments?)

So what would linguistic theorizing look like if efforts were made to address some of the issues raised in this section? Here is not the place for the details, some of which are substantial. However, we can get a partial glimpse of what research will look like when aggregative methods are made more systematic. In the simplest situation, we might summarize how Hornstein will defend his theory as superior to Landau’s as follows:

Taking the evidence from the previously cited sources, as well as data new to this paper, H used method  $M_1$  to assess the relative degrees of independence of the total number  $k$  of sentences used in this study. After controlling for the various dependencies, the data set had an estimated size of  $n$ , and was distributed as follows:  $n_1$  of them were correctly predicted (by my model) to be grammatical,  $n_2$  of them were incorrectly predicted to be grammatical,  $n_3$  were correctly predicted to be ungrammatical, and  $n_4$  were incorrectly predicted to be ungrammatical. H then used method  $M_2$  with this array to arrive at the raw coefficient  $c$  of the degree of association between his theory and the empirical facts. Method  $M_3$  was then used to adjust  $c$  to account for the size  $n$  of the data set and the degree of simplicity of the model. The result was a measure  $m$  of the theory. This same procedure was then performed on L’s theory, using the available

appropriate data, arriving at another measure  $m'$ . Since the ratio  $m/m'$  exceeded the threshold  $t$ , this analysis supports H's theory over L's.

(Elsewhere (Johnson ms.) I have developed and defended the structure of this particular procedure. E.g., the nature of linguistic theorizing makes it natural to regard  $n_1 - n_4$  as Poisson distributed random variables, and that a natural value of  $M_1$  is a (possibly weighted) phi coefficient. Moreover, the nature of linguistic inquiry makes the batch-processing methods of classical statistics inappropriate; instead, a Bayesian or likelihood approach (e.g., Lee 2004, Royall 1997, Blume 2002) which allows for continuous updating of data is more suitable.)

There are many advantages to pursuing an aggregative method such as the one sketched above, assuming  $M_1 - M_3$  are mathematical algorithms of some form. I will list just three, all of which concern the explicitness of the model (in contrast to the method of relying on informal subjective professional judgments). First, the approach makes its assessment of the relative evidential strength for one theory over another perfectly clear. That is, it both indicates precisely *how* strong the evidence is for a given theory, and it indicates *how* strong the support (according to the method) is for one theory over another. Moreover, by considering the standard deviations and (what corresponds to) confidence intervals of these statistics, one can precisely estimate the degree of accuracy of this assessment. These assessments may need fine-tuning, or may be wrong, but there is no doubt what they are or how they are related to the data and theoretical considerations. In general, the explicitness of the model makes it easier to identify and fix flaws. The approach also makes explicit exactly what role each consideration or datum is doing, and how important it is. When humans try to 'get a feel' for the general trend of the evidence, experiments show that our interpretations are often biased towards what we happen to antecedently believe or want to be true. This does not happen within the present approach.

Second, the model also helps to identify and emphasize the sources of indeterminateness in linguistic theorizing. Linguists commonly assume that they are studying fully determinate, mathematically precise algorithms – aka ‘grammars’. I have adopted this assumption, too. Thus, a grammar and its set of possible outputs (i.e., expressions of the language) are discrete, fully determinate mathematical objects with no stochastic features. But we’ve seen that there nevertheless are probabilistic elements in the *process* of linguistic inquiry into these grammars. One source of stochasticity is the radical incompleteness of actual present-day linguistic theories. Thus, in actual research, we have to assess the probability that a current theory can be developed so as to appropriately handle whatever challenges it faces. This radical incompleteness of linguistic theories explains why a given theory is not instantly refuted by a single ‘counterexample’. In practice, counterexamples don’t really address a particular existing theory; rather they raise questions about the likelihood of finding a theory that not only contains the crucial aspects of the current one, but also accounts for the counterexample in a sufficiently scientifically satisfying manner. Another source of stochasticity comes from our incomplete access to the empirical data. Since we don’t know what relevant expressions some clever linguist will discover in the future, at present we have to assess the probability that a current theory will continue to be good in light of new data. (The next paragraph displays a third type of stochasticity.)

Third, the approach helps identify various background assumptions in play. For instance, the procedure begins by estimating the degree of independence between the various presented data sentences. This crucial step explicitly addresses a problem that every linguist faces, namely that of estimating how much data there is in the first place. After all, if one claims that, say, *Mary wants to win* supports a given theory about control, one can’t maintain that *Sue wants to*

*win* is another, independent, bit of evidence. The two sentences are simply too similar in the relevant respects. Indeed, for present purposes, the two sentences should count as only one datum. But given a particular project, how closely related various sentences are to one another is a matter of degree. Two sentences that are somewhat, but not entirely, related to one another (given the interests of the particular project) should be counted as more than one but less than two distinct data points. E.g., *try* and *prefer* behave differently, but not always, so how much data is expressed in Hornstein's (2003, 11) *John tried/preferred PRO eating a bagel?*

Evidentially speaking, does this count as one sentence? Two? Somewhere in between? Surely we have a right to ask how much data one has for one's theory. Estimating and aggregating these associations across a normal-sized data set is not something that can be done reliably by informal (usually tacit) judgment. However, such estimations can be calculated by algorithms which themselves can be studied, criticized, and improved.

It must be stressed that the precise methods of statistics and mathematical modeling do not supply a methodological panacea, especially in the social sciences. Quantitative methods and their results are well known to be susceptible to misuse and misinterpretation. Furthermore, such methods frequently use only limited types of information, and many considerations that might be relevant are not factored into the model. Consequently, such models often support limited interpretations that are considerably weaker than the 'big picture' theories they are used to support. But as we've just seen, this infamy reflects the real strength of formal methods. Their consistent and explicit structure, in contrast to informal, subjective professional judgments, often makes it relatively easy to uncontroversially identify misuses or overinterpretations. Similarly, one can assess what advantages are gained by adjusting the method, perhaps to include new types of information. And even when the guidance these methods supply is less than total, which

it often is, the mathematical ‘guidrails’ they do supply can be of great value (cf. Trout 1998). If the history of the other sciences is any guide, the greatest advantage of such methods is not so much to answer the big questions, but to tightly constrain the room that researchers have for speculation and informal hypothesizing. Indeed, when used correctly, such techniques often answer questions and quell disputes before they arise.

## 4 Conclusion

The really important aspects of scientific methodology are subtle, and they tend to be encoded quantitatively. On the one hand, linguistic methods realize some of these aspects non-quantitatively, and many features of linguistic theorizing are thereby rendered scientifically legitimate. Because linguistic models are only partial, and thus have significant (non-quantitative) ‘error terms’, the Problem of Unsaved Phenomena disappears. Similarly, because hypotheses about unobserved linguistic structures are constrained by the interaction of multiple (non-quantitative) measurements, the Problem of Undefined Terms is eliminated. On the other hand, linguistics is distinctive among the sciences in that it does not use mathematical methods to mediate the relationships between evidence and theories. Instead, various estimations must be made by informal, subjective professional judgments. Such inferential strategies are notoriously unreliable in general. The complexity of linguistic theories – coupled with the marked lack of feedback available to hone such judgments – makes such strategies especially worrisome. In short, methods for manipulating evidence in linguistic theorizing presents an important but neglected area of linguistic research. Our methods have not ‘exhausted the methods of science’ (Chomsky 1986, 252); rather, they have only begun to scratch the surface.

*Department of Logic and Philosophy of Science*

*University of California, Irvine*

## References

- Allen, K. 2003: Linguistic metatheory. *Language Sciences*, 25, 533 – 560.
- Baker, M. 1988: *Incorporation*, Chicago: University of Chicago Press.
- Basilevsky, A. 1994: *Statistical Factor Analysis and Related Methods*. New York: Wiley-Interscience.
- Batchelder, W. and Romney A. 1988: Test theory without an answer key. *Psychometrika*, 53, 71 – 92.
- Bishop, M. and Trout, J. 2002: 50 years of successful predictive modeling should be enough: lessons for philosophy of science. *Philosophy of Science*, 69, S197 – S208.
- Bishop, M. and Trout, J. 2005: *Epistemology and the Psychology of Human Judgment*. Oxford: OUP.
- Blume, J. 2002: Likelihood methods for measuring statistical evidence. *Statistics in Medicine*, 21, 2563 – 2599.
- Brody, M. 1999: Relating syntactic elements: remarks on Norbert Hornstein's 'Movement and chains'. *Syntax*, 2, 210 – 226.
- Burnham, K. and Anderson, D. 2002: *Model Selection and Multimodel Inference* (2nd ed.). New York: Springer.
- Cappelen, H. and Lepore, E. 2005: *Insensitive Semantics*. Oxford: Blackwell's.
- Carston, R. 2004: Relevance theory and the saying/implicating distinction. In L. Horn and G.

- Ward (eds.) *Handbook of Pragmatics*. Oxford: Blackwell.
- Chomsky, N. 1965: *Aspects of the Theory of Syntax*, Cambridge, MA: MIT Press.
- Chomsky, N. 1975: *Reflections on Language*. New York: Pantheon.
- Chomsky, N. 1986: *Knowledge of Language*, Westport, Conn.: Praeger.
- Chomsky, N. 1988: *Language and Problems of Knowledge*. Cambridge, MA: MIT Press.
- Chomsky, N. 1992: Explaining language use. *Philosophical Topics*, 20, 205 – 231.
- Chomsky, N. 1995: *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, N. 2000: *New Directions in the Study of Language and Mind*. Cambridge: CUP.
- Chomsky, N. and Lasnik, H. 1993: The theory of principles and parameters. Reprinted in  
Chomsky 1995.
- Christie, Katrien N., and Phillip Christie, 1999: Gambling on UG: the application of Monte Carlo  
computer simulation to the analysis of L2 Reflexives. *Syntax* 2:2 80 – 100.
- Culicover, P. 1997: *Principles and Parameters*. Oxford: OUP.
- Culicover, P. and Jackendoff, R. 2001: Control is not movement. *Linguistic Inquiry*, 32, 493 –  
512.
- Davidson, D. 1986: A nice derangement of epitaphs. In E. Lepore (ed.), *Truth and  
Interpretation: Perspectives on the Philosophy of Donald Davidson*. Oxford:  
Blackwell's.
- Dawes, R. 1979: The robust beauty of improper linear models in decision making. *American  
Psychologist*, 34, 571 – 582.
- Dowty, D. 1991: Thematic proto-roles and argument selection. *Language*, 67, 547 – 619.
- Faust, David 1984: *The Limits of Scientific Reasoning*. Minneapolis, University of Minnesota  
Press.

- Fodor, J. 1998: *Concepts*, Oxford: Clarendon.
- Fodor, J. and Lepore, E. 2005: Impossible words: a reply to Kent Johnson. *Mind and Language*, 20, 353 – 356.
- Forster, M. and Sober, E. 1994: How to tell when simpler, more unified, or less ad hoc theories will provide more accurate predictions. *British Journal for the Philosophy of Science*, 45, 1 – 35.
- Grice, H. and Strawson, P. 1956: In defense of a dogma. *The Philosophical Review*, 65, 141 – 158.
- Grimshaw, J. 1990: *Argument Structure*. Cambridge, MIT Press.
- Grove, W. and Meehl, P. 1996: Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: the clinical-statistical controversy. *Psychology, Public Policy, and Law*, 2, 293 – 323.
- Haegeman, L. 1994: *Introduction to Government and Binding Theory* (2nd ed.) Oxford: Blackwell.
- Hale, K. and Keyser, S. 1986: Some transitivity alternations in English. Lexicon Project Working Papers 7, Cambridge, MA: Center for Cognitive Science, MIT.
- Hale, K. and Keyser, S. 1987: A view from the middle. Lexicon Project Working Papers 10, Cambridge, MA: Center for Cognitive Science, MIT.
- Hale, K. and Keyser, S. 1992: The syntactic character of thematic structure. In I. Roca (ed.), *Thematic Structure and its Role in Grammar*. New York: Foris.
- Hale, K. and Keyser, S. 1993: On argument structure and the lexical expression of syntactic relations. In K. Hale and S. Keyser (eds.), *The View from Building 20*. Cambridge, MA: MIT Press.

- Hale, K. and Keyser, S. 1999: A response to Fodor and Lepore, 'Impossible words?'. *Linguistic Inquiry*, 30, 453-66.
- Hale, K. and Keyser, S. 2003: *Prolegomenon to a Theory of Argument Structure*. Cambridge, MA: MIT Press.
- Henrion, M. and Fischhoff, B. 1986: Assessing uncertainty in physical constants. *American Journal of Physics*, 54, 791 – 798.
- Hogg, Robert V., and Allen T. Craig 1995: *Introduction to Mathematical Statistics*. New York: Macmillan
- Hornstein, N. 1998: Movement and chains. *Syntax*, 1, 99 – 127.
- Hornstein, N. 1999: Movement and control. *Linguistic Inquiry*, 30, 69-96.
- Hornstein, N. 2000: On A-chains: a reply to Brody. *Syntax*, 3, 129 – 143.
- Hornstein, N. 2003: On control. In R. Hendrick (ed.), *Minimalist Syntax*. Oxford: Blackwell.
- Jackendoff, R. 1983: *Semantics and Cognition*, Cambridge, MA: MIT Press.
- Jackendoff, R. 1987: The status of thematic relations in linguistic theory. *Linguistic Inquiry*, 18, 369-411.
- Jackendoff, R. 1990: *Semantic Structures*, Cambridge, MA: MIT Press.
- Jackendoff, R. 1997: *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.
- Jackendoff, R. 2002: *Foundations of Language*. Oxford: OUP.
- Johnson, K. 2006: [Forthcoming]: Tacit and accessible understanding of language. *Synthese*,
- Johnson, K. ms. 'Analytic Methods for Linguistics'
- Kahneman, D. and Tversky, A. 1984: Choices, values, and frames. *American Psychologist*, 39, 341 – 350. Reprinted in Kahneman and Tversky 2000, 1 – 16.
- Kahneman, D. and Tversky, A. 2000: *Choices, Values, and Frames*. Cambridge: CUP.

- Landau, I. 2000: *Elements of Control: Structure and Meaning in Infinitival Constructions*.  
Dordrecht: Kluwer.
- Landau, I. 2003: Movement out of control. *Linguistic Inquiry*, 34, 471 – 498.
- Lee, Peter M. 2004: *Bayesian Statistics* (3<sup>rd</sup> ed.). New York: Hodder Arnold.
- Leamer, E. 1985: Sensitivity analyses would help. *American Economic Review*, 75, 308 – 313.
- Levin, B. and Rappaport Hovav, M. 1995: *Unaccusativity: At the Syntax--Lexical Semantics Interface*, Cambridge, MA: MIT Press.
- Liu, C. 2004: Laws and models in a theory of idealization. *Synthese*, 138, 363 – 385.
- Malinowski, E. 2002: *Factor Analysis in Chemistry*. New York: John Wiley and Sons.
- Manzini, R. and Roussou, A. 2000: A minimalist theory of A-movement and control. *Lingua*, 110, 409 – 447.
- Meehl, Paul E, 1986: Causes and effects of my disturbing little book. *Journal of Personality Assessment* 50:3, 370 – 375.
- Meehl, P. 1954: *Clinical versus Statistical Prediction: A Theoretical Analysis and Review of the Evidence*. Minneapolis: University of Minnesota Press.
- Nickerson, Raymond S. 1998. “Confirmation Bias: A Ubiquitous Phenomenon in Many Guises”.  
*Review of General Psychology* 2:2, 175 – 220.
- Parsons, T. 1990: *Events in the Semantics of English*, Cambridge, MA: MIT Press.
- Pesetsky, D. 1995: *Zero Syntax*. Cambridge, MA: MIT Press.
- Pietroski, P. 2005: *Events and Semantic Architecture*. Oxford: OUP.
- Prince, A. 2002a: Fundamental properties of harmonic bounding. Ms. Rutgers University.  
[http://rucss.rutgers.edu/tech\\_rpt/harmonicbounding.pdf](http://rucss.rutgers.edu/tech_rpt/harmonicbounding.pdf)
- Prince, A. 2002b: Arguing optimality. Ms. Rutgers University.

<http://ling.rutgers.edu/gamma/argopt.pdf>

- Putnam, R. 2000: *Bowling Alone* New York: Touchstone.
- Recanati, F. 2001: What is said. *Synthese*, 128, 75 – 91.
- Royall, R. 1997: *Statistical Evidence: A Likelihood Paradigm*. Boca Raton, FL: Chapman and Hall.
- Schütze, C. 1996: *The Empirical Base of Linguistics*. Chicago: University of Chicago Press.
- Stone, T. and Davies, M. 2002: Chomsky amongst the philosophers. *Mind and Language*, 17, 276 – 289.
- Swets, J. 1996: *Signal Detection Theory and ROC Analysis in Psychology and Diagnostics: Collected Papers*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Swets, J., Dawes, R. and Monahan, J. 2000: Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest*, 1, 1 – 26.
- Townsend, D. and Bever, T. 2001: *Sentence comprehension*. Cambridge, MA: MIT Press.
- Trout, J. D. 1998: *Measuring the Intentional World*. Oxford: OUP.
- Tversky, A. and Kahneman, D. 1971: Belief in the law of small numbers. *Psychological Bulletin*, 76, 105 – 110.
- Walenski, M. and Ullman, M. 2005: The science of language. *The Linguistic Review*, 22, 327 – 346.
- Whewell, W. 1840: *The Philosophy of the Inductive Sciences*. London.
- Williams, E. 1980: Predication. *Linguistic Inquiry*, 11, 203 – 238.
- Williams, E. 1983: Against small clauses. *Linguistic Inquiry*, 14, 287 – 308.
- Wolff, Phillip 2003: Direct causation in the linguistic coding and individuation of causal events. *Cognition* 88, 1 – 48.

Zucchini, W. 2000: An introduction to model selection. *Journal of Mathematical Psychology*,  
44, 41 – 61.