Notational Variants and Invariance in Linguistics*

(Forthcoming in *Mind and Language*)

Abstract. This paper argues that the much-maligned 'notational variants' of a given formal linguistic theory play a role similar to alternative numerical measurement scales. Thus, they can be used to identify the invariant components of the grammar; i.e., those features that do not depend on the choice of empirically equivalent representation. Treating these elements as the 'meaningful' structure of language has numerous consequences for the philosophy of science and linguistics. I offer several such examples of how linguistic theorizing can profit from adopting a measurement-theoretic viewpoint. The first concerns a measurement-theoretic response to a famous criticism of Quine's. Others follow from issues of simplicity in the current biolinguistics program. An unexpected similarity with behaviorist practices is also uncovered. I then argue that managable and useful steps can be taken in this area.

1. Introduction

* Thanks are due to Bob Matthews, who offered numerous comments on an earlier draft of this paper. More generally, he taught me the value of extending the concepts of classical measurement theory to non-numeric representational domains. The central ideas of the present paper owe a great deal to his work on this topic; e.g., in his 1993, 2007. Also, an anonymous reviewer for this journal offered a great deal of thoughtful and helpful feedback, which substantially improved this paper.
Address for correspondence: Kent Johnson, Department of Logic and Philosophy of Science, 3151 Social Science Plaza A, University of California, Irvine, Irvine, CA 92697-5100
Email: johnsonk@uci.edu

In linguistics, one sometimes encounters the claim that one theory is a 'notational variant' of another. Such claims are often used to criticize one theory as not differing empirically from some older, more familiar one.[1] One of the earliest uses of the term is in Chomsky's discussion of Katz and Postal's theory of generative semantics (originally published in 1970, circulated as a pamphlet in 1968; cf. Moultin and Robinson, 1981, 229, fn. 9). Chomsky writes:

> Given alternative formulations of a theory of grammar, one must first seek to determine how they differ in their empirical consequences, and then try to find ways to compare them in the area of difference. It is easy to be misled into assuming that differently formulated theories actually do differ in empirical consequences, when in fact they are intertranslatable – in a sense, mere *notational variants* (Chomsky, 1972, 69, emphasis added; cf. also Fromkin, 2000, 705, Moultin and Robinson, 1981, 229–232).

More recently, a notational variant was characterized as:

> A notational model (for example, a model of a theory of grammar) that represents the same set of abstract properties as an alternative notational model, but in a superficially different way, and which makes the same empirical predictions as the alternative model. (The relation of notational variance is symmetrical – that is, if X is a notational variant of Y, than Y is a notational variant of X.) (Fromkin, 2000, 705).

The idea here is simple. Two theories may *prima facie* appear drastically different, and yet be indiscernable in terms of their empirical predictions, etc. In such a case, they are not essentially different and may be assumed to each capture one and the same underlying idea, albeit in distinct vocabularies. But then, aside from practical or aesthetic preferences, there is no

[1] The label is also sometimes used to show how a new theory can co-opt some advantages of another theory, while avoiding some of the latter's disadvantages. Assuming the disadvantages in question have real empirical bite, these theories would not here count as genuine notational variants.

reason to favor one over the other. For this reason, linguists sometimes explicitly defend their favored theory against the (anticipated) charge that it is simply a notational variant of some more familiar view; e.g., Moultin and Robinson, 1981, 229– 232. On its face, then, the charge of being a notational variant would seem to be a real criticism of a theory, and often it is. Researchers usually don't intend to simply recast old theories in new ways; the idea is rather to find some characterization that better captures (explains, describes, predicts, etc.) the phenomena under study than the rival theories. But could notational variants ever serve some purpose; could they be *important*? The short answer is Yes.

A word on the notion of a 'notational variant'. As the term is used here, two theories (formal grammars, etc.) are notational variants iff they are empirically equivalent, in the sense that, borrowing from Chomsky's quote above, 'they do not differ in their empirical consequences'. In this sense, to be empirically equivalent, two theories must be equivalent in terms of *all* their empirical consequences, not just the obvious or intended ones. In particular, there is no suggestion that determining a theory's empirical consequences is a simple matter – a fact that looms large in this essay. More generally, whether two theories are notational variants is not an epistemic notion at all. Although nobody owns the rights to this term, my characterization fits quite well with its ordinary usage. After all, the charge that one theory may be a notational variant of another typically doesn't imply that their equivalence is obvious in any sense – if it were, then surely no bright linguist(s) would waste time developing such counterparts to currently existing theories. Similarly, it wouldn't be 'easy to be misled' (cf. Chomsky above) into thinking two notational variants are substantively different. More importantly, the 'obviousness' of two theories' equivalence may differ across researchers, but grammar A shouldn't be a notational variant of grammar B for you, but not for me.

This essay argues that the notational variants of a theory constitute an important part of linguistic theorizing, and urges that they should be a more central component of linguistic research. Currently, there is little explicit discussion of these issues; but I will argue that they are crucial to understanding the content of a theory. Collectively, the notational variants of a theory determine the empirically 'real' or 'meaningful' structure of any one of the theories taken individually. This meaningful structure is often not identifiable without recourse to notational variants (i.e., symmetries). Moreover, even when all the notational variants cannot be articulated, individual members can nonetheless serve to partially distinguish a theory's real empirical content and its merely artifactual additional structure.

Thus, my positive proposal is methodological: a measurement-theoretic perspective can and should be pursued in generative linguistics. This proposal is as important to linguistics as it is to many or most other areas of scientific inquiry, since it involves, inter alia, the careful separation of those aspects of the formal structure of a model that correspond to the empirical phenomenon in question, and those that do not. (In this sense, it's a bit surprising that measurement-theoretic considerations haven't already assumed a prominent role in theoretical linguistics, whose stock in trade is exceedingly complex and subtle formal models of the faculty of human language

.) In support of this thesis, I will defend several further subtheses, most prominently:

1. Quine's (1972) argument against a grammar's being treated as guiding human linguistic abilities is exactly backwards. The existence of (possibly infinitely) many empirically equivalent grammars is not a *problem* for understanding the true mechanisms underlying language; rather, it's *crucial* for such understanding.

2. A measurement-theoretic notion of meaninglessness is *necessary* for any notion of the 'simplicity' of a grammar, and this is all the more so for the kind of simplicity employed in the current biolinguistics program.
3. Biolinguists' frequent talk of 'virtually conceptually necessary' structure should be understood carefully, since in practice the articulation of any such structure may require additional, empirically meaningless structure. In a slogan, it's probably conceptually necessary that the conceptually necessary bits of grammar require conceptually irrelevant elements!
4. Interestingly, a measurement-theoretic perspective shows that some key features of contemporary biolinguistics and behaviorist psychology are closer than one might have expected.
5. Even a simplified version of a very simple linguistic process like Merge has some nontrivial features; but we can see how to begin some measurement-theoretic work on these key components of grammar.

The paper is structured as follows. §2 spells out the basic measurement-theoretic perspective for linguistic theorizing. The idea of using measurement-theoretic concepts to investigate high-level issues in the philosophy of psychology is due to Robert Matthews (Matthews, 1994, Matthews, 2007) (cf. e.g., Stalnaker, 1984, Churchland, 1979, Field, 1981, Davidson, 1989 for some earlier gestures); my use of this idea here is very different from his. §3 turns to Quine's famous charge that grammars can at best 'fit' a speaker, and may never be said to 'guide' her. The first subthesis listed above is defended here. §4 considers some contemporary issues of the 'simplicity' of grammatical processes, particularly as they arise for biolinguistics. The next three subtheses, concerning simplicity, conceptually necessary features, and behaviorism, are defended in this section. §5 gestures at some ways that measurement-theoretic progress can be had, even in the face of the substantial uncertainties that surround hypothesized (partial) grammars of FHL. An example is given there that establishes the final subthesis. §6 concludes the paper.

In what follows, I adopt a familiar view of linguistics (cf. e.g., Chomsky, 1986, 25–26 for a succinct summary). According to this view, at some scientifically interesting level(s) of (psychological, neurological, biological) characterization, there exists a 'mental organ', the *faculty of human language*, FHL, that is responsible for normal humans' (in normal environments) abilities to acquire and speak a language (where this latter term is used somewhat loosely here; cf. Chomsky, 1995b, 1; cf. Collins, 2004 for discussion). As with other unobserved hypothetical entities, the existence and nature of FHL is justified by the overall fruitfulness of this theoretical posit in a developing broader psychological theory. In a normal developmental environment, a child's FHL, the hypothesis goes, attains some particular state, an *I-language*. The latter algorithmically produces certain kinds of cognitive data structures, *I-expressions*, which don't necessarily correspond directly with any pre-theoretic notion of linguistic expressions. Rather, I-expressions are assumed to interact with various other components of a person's psychological (and neurological, biological, etc.) makeup so as to ultimately result in key aspects of the comprehension and production of natural language. Nevertheless, the hypothesis goes, there is scientific merit in individuating something like FHL that plays something like the central role just described. Figuring out the nature of FHL and just what kind of role it plays is the central goal of linguistics. Thus, theories about FHL along with their notational variants, is the topic in what follows. (For stylistic convenience, I will frequently focus on grammars below; however corresponding considerations apply to other components of FHL.)

2 Representation and Invariance in the Measurement of Temperature and Language

This section briefly introduces the basic structure of a measurement-theoretic analysis of an empirical phenomenon. I use a (part of) the analysis of temperature as an example.[2] For an excellent historical study of the deeper issues for temperature and thermometry, cf. Chang, 2004.

Generally speaking, measurement theory is the mathematical study of how mathematical representations, typically numbers, are associated with various types of empirical phenomena Krantz et al., 1971, Suppes et al., 1989, Luce et al., 1990. The subject matters are formal models of some interesting bit of empirical phenomena, characterized by some set of empirically relevant axioms. One goal of classical measurement theory is to prove a *representation theorem*, showing that the empirical model's structure can be represented numerically, in a structure-preserving way. Often there is more than one such numerical representation. Thus, a second related goal is to prove a *uniqueness theorem*, which specifies the set of numerical representations that satisfy the representation theorem.

For example, temperature ranges – i.e., pairs of temperatures – are assumed to be ordered, in that the relation ab □ cd, meaning 'temperature range ab is at least as great as temperature range bc' is transitive and total over temperature ranges. This relation, □, and the temperature ranges are assumed to satisfy further axioms, such as 'If ab □ xy and bc □ yz, then ac □ xz'; cf. Krantz et al., 1971, chap. 4, defs. 1, 3 for details. With these assumptions explicitly in place, a representation theorem can be proved:

[2] More specifically, the discussion below concerns temperature characterized as an algebraic-difference structures; for complete details, cf. Krantz et al., 1971, Chap. 4, definition 3 and theorem 2.

(1) There exists a homomorphism f from the qualitative temperature structure to the additive reals such that ab □ cd iff $f(b) - f(a) \geq f(d) - f(c)$.

Similarly, a uniqueness theorem is also provable:

(2) A function g also satisfies (1) iff $g = \alpha f + \beta$, where $\alpha > 0$.

(f, of course, is a g: $f = (1)f + 0$.)

(1) shows that it's possible to represent temperatures numerically, so that numerical differences really do represent temperature intervals, and the ordering of the real numbers represents the intervals' ordering according to size. (2) shows that there are in fact many ways of constructing such a representation. In terms of satisfying the criteria in (1), any of them will do; they are *notational variants*. In other words, the differences between such representations are of no empirical consequence. Crucially for present purposes, this means that a (numeric) feature of one representation that is not shared by *all* the others – such as which temperature is assigned the value 0 – does not represent a feature in the empirical structure being represented. Such a feature is, by definition, empirically *meaningless*, a mere artifact of a particular representation. On the other hand, a necessary and sufficient condition for a feature of a given representation to be empirically *meaningful* is that the feature be present in every representation that satisfies (1).[3] This criterion of empirical meaningfulness, invariance across all the adequate representations, is obviously satisfied by ratios of intervals: $[f(a)-f(b)]/[f(c)-f(d)] = [g(a)-g(b)]/[g(c)-g(d)]$.

---

[3] There are several characterizations of empirical meaningfulness in the literature; when made precise, many of them are equivalent to the one used here (reference invariance), which is the

Similarly, it is not satisfied by ratios of individual temperature representations: f(a)/f(b) will not in general equal g(a)/g(b). Moreover, this kind of invariance is a fundamental concept across the sciences. E.g., the mathematical psychologists William Batchelder and Louis Narens write:

> There is an old saw in mathematics that most things of mathematical or theoretical interest are invariant under important classes of transformations. For example, topological properties of a rubber sheet are invariant under various distortion operations such as stretching, and the areas and volumes of geometrical objects are invariant under the transformation of rotating coordinate system, etc. Invariants have also played a major role in the development of science, e.g., the conserved quantities of mass, energy, and momentum are invariant under radically different description frameworks. Even social science has its invariants, e.g., the correlation coefficient as well as the t and F statistics are invariant under linear transformations in the dependent variables. Batchelder and Narens, 1977, 114; cf. Narens, 2007, 41 for similar sentiments stemming from Felix Klein's Erlanger program in geometry.

Given what we know about the temperatures, any g is a fully acceptable characterization. Superficially, however, they might appear vastly different. Even in the very simplest cases (cf. Suppes, 2002), f could leap in huge bounds across the negative real numbers, perhaps yielding only irrational values, whereas g might yield only positive integers, or only rational values compressed into some tiny region between, say 151.3 and 151.36.[4] Nevertheless, they have the 'same empirical consequences', and are 'intertranslatable' (Chomsky, 1972, 69). In particular they each employ some different formal structure which can be systematically exchanged, transforming the one into the other.

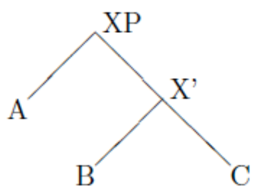
---

most applicable for linguistics; cf. Luce et al., 1990, Narens, 2002, Narens, 2007 for technical discussion; Dresner, 2010, Dresner, 2004 for relevant philosophical discussion.

[4] Indeed, as Krantz, et al. discuss (Krantz et al., 1971, 152), matters could be even more extreme than this. We might further transform g to, say, an exponential function $g'(a)= e^{g(a),}$ so that ab □ cd iff g'(a)/g'(b) ≥ g'(c)/g'(d: But a similar argument shows that f and g' are still mere notational variants in the relevant sense.

Abstractly, the transition from temperature to linguistics is straightforward. Where a particular temperature scale represents temperatures with numbers, a grammar represents linguistic phenomena with formal characterizations of constituent structure, along the lines illustrated in (3):

(3) a. $[_{XP}$ A $[_{X'}$ B C]

b. 

Obviously, representations like (3) lack many numerical properties; however, they nevertheless play the same role as numeric representations of temperature. Similarly, the representation of individual elements – whether particular temperatures by particular numbers, or particular expressions by particular formal structures – are themselves embedded in larger representations, namely particular whole temperature scales[5] and particular grammars. The notational variants of a given grammar are those grammars that map expressions to explicit formal representations that similarly respect the empirical information about the expressions (more on this below).

(It should be noted at the outset that there is nothing untoward about employing the basic ideas of measurement theory in a non-numeric context. An examination of the details of the proofs of the representation and invariance of colors (and, interestingly, force) in the highly-respected *Foundations of Measurement* shows that the representing (infinite-dimensional) vectors can themselves be literally built out of the elements of the empirical structure, which are themselves, intuitively, colors Suppes et al., 1989, chap. 15. Similarly, although it is not typically thought of as a measurement-theoretic matter, the (in)famous issue of 'factor indeterminacy' in fact shows that the representation of latent empirical structure by random variables is unique only up to some further arbitrary constituent random variable satisfying certain rather weak conditions Guttman, 1955, Maraun, 1996.)

Even a very simple example like (3) displays some of the relevant issues concerning empirical meaningfulness. E.g., the familiar linear and tree-structure representations of constituent structure are empirically equivalent (as standardly interpreted). Thus, in (3b), no linguistic meaning is attached to the length of the lines on the tree, or the size of the angles between them; only certain ordering properties induced by the tree are relevant; e.g., Berwick et al., 2011, 11–13, Chomsky, 2005, ff. Similar remarks apply for the spacing between the brackets in (3a), or the use of square instead of round brackets, subscripts instead of superscripts, etc. Importantly, the empirically relevant structure is contained in both (3a-b), and any features not shared by both are empirically meaningless representational artifacts. Indeed, the empirically relevant structure is *exactly* what is present in all possible notational variants of the representations in (3).[6]

Table 1 sums up the main parallels between temperature measurement and formal grammars.

---

[5] The term *scale* is often used in the technical literature to refer to a broader class of representations; here, though, I use it in the familiar sense in which we speak of the Fahrenheit or Celsius scales.

[6] Actually, (3) illustrates two distinct types of notational variance. On the one hand, (3a-b) can be understood as each referring to one and the same underlying structural/mathematical object, but in different terminology. In this case, the equivalence of (3a-b) is trivially ensured. Alternatively, (3a-b) could also be interpreted as referring to distinct mathematical objects, one geometrical and the other sequential, perhaps. This case is more akin to the temperature example, where distinct mathematical objects are used to represent one and the same empirical phenomenon. In this case, the empirical equivalence of (3a-b) is not guaranteed, but instead depends on how the

Feature	Temperature	Language
System representation	$f : \mathcal{T} \to \mathbb{R}$	Formal grammar
Individual representations	Real numbers	Individual representations of I-language expressions (indexed by I-language)
System-level variable features	location, scale parameters	Peculiar aspects of notational variants of formal grammars of FHL
Individual-level variable features	numeric values of particular temperatures	Peculiar aspects of individual representations of I-expressions, of angle, arm length, bracket choice of individual representations, etc.
True structure	affine structure of temperature	"linguistic reality", including (partly) non-empirical structure
Defining empirical constraints	Empirical axioms	Empirical linguistic data points, generalizations

Table 1: Summary of parallel structure between the measurement of temperature and the formal grammatical representation of FHL.

From the current perspective, some further measurement-theoretic features of linguistic theorizing can be identified. First, if the empirical facts a grammar of FHL is responsible for are held fixed, then these facts induce a partition of all possible grammars into equivalence classes of notational variants. Clearly, the binary relationship X is a *notational variant* of Y is symmetric. But it is also reflexive (every theory is trivially a notational variant of itself) and transitive (if X can be notationally translated into Y, and Y into Z, then X can be translated into Z). Thus, this relation supplies the relevant partition. Moreover, two grammars of FHL make empirically

---

mathematical objects they denote are themselves related to the structure of the empirical object distinct predictions if and only if they belong to distinct elements of this partition, and they are notational variants iff they belong to the same one. Importantly, the empirical consequences, predictions, etc. are to be given explicitly in the empirical axioms. Thus, all sorts of factors that might otherwise not be considered consequences of a theory (e.g. Laudan and Leplin, 1991) would here uncontroversially count as features that could be axiomatized and thereby fall under the present notion of an 'empirical prediction', consequence, etc. In fact, measurement-theoretic approaches can often act as a kind of neutral format where constraints from all over – cross-linguistic facts, issues of learnability, constraints on integration of FHL with larger psychological theory, etc. – are collected together and imposed simultaneously. (Indeed, there is no principled reason why non-empirical considerations might not be included as well.) Thus, if we individuate a theory by its empirical consequences independently from how we may choose to state it, a theory of FHL can be identified with its entire equivalence class of notational variants, not with any one of them.[7] For clarity, terms like 'theories', 'grammars' etc. below will be used in the more specific sense. Thus, by stipulation, theories and grammars can have notational variants; they are themselves not the equivalence classes just mentioned. (All the points to be made below could easily be translated into points about theories, etc. understood as equivalence classes.)

Secondly, there is further structure *within* each class of notational variants. Since these grammars are 'intertranslatable' (Chomsky, 1972, 69), there is a collection of transformations that turn a grammar X into its various notational variants. This collection of transformations forms a mathematical group under the ordinary composition of transformations.[8] This is easy to see. The

---

in question.

[7] Alexander George offers five useful distinctions in his discussion of how not to become confused about linguistics George, 1989. We might add to these the distinction between a particular notation variant and its entire equivalence class, to further reduce confusion.

[8] A group is any set A of objects with a binary function $*$: $A \times A \to A$ such that (i) $*$ is associative: $a * (b * c) = (a * b) * c$; (ii) A contains an identity element e: $a * e = e * a = a$; and (iii) every element a of A has an inverse a^{-1} in A: $a * a^{-1} = a^{-1} * a = e$. In the text above,

null transformation e that transforms every X into itself is the identity element. By the symmetry of the notational variant relation, a transformation t that turns some Xs into some Ys also implicitly defines a means for turning the Ys back into the original Xs. Thus, each transformation has an inverse t^{-1}.[9] Similarly, combining such transformations is obviously an associative operation. The transformation group of a collection of objects is often key to understanding the collection's internal structure, particularly concerning issues of invariance, and hence empirical meaningfulness. Moreover, adding more empirical constraints – axioms in the standard measurement-theoretic cases; data, generalizations, etc. in the case of linguistics – produces a subgroup of the original group. (E.g., (2) shows that for temperature, the transformation group is the set of positive affine functors: $\{T : T(f)= \alpha f + \beta, \alpha > 0\}$. If additional information is added about, say, where the true zero-point is, the new transformation group would be the subgroup of positive linear functors $\{T : T(f)= \alpha f, \alpha > 0\}$.) Thus, if an 'interesting' transformation group can be found using only some of the empirical facts about FHL, then any further empirical embellishments will result in a transformation group contained in the former one.

This section has presented the basic measurement-theoretic picture. As mentioned above, my own view is that this perspective should be developed and pursued in more detail in generative linguistics. In the following sections, I motivate this in three distinct ways. The present perspective can resolve old controversies (§3), clarify and better orient some current views (§4), and, pragmatically speaking, it can be employed at present, in partial and developing fashion (§5).

---

it is the *transformations* that constitute the objects a, b, c, etc. of the group; $*$ composes these transformations so that for any grammar X, $a * b(X)= a(b(X))$.

[9] It can seem awkward to think of a means of transforming one theory X into another theory Y as a total function whose domain includes all theories X in the equivalence class. However, there are no real difficulties here. Transformations – i.e., means by which the notation of X is transformed into the notation of Y – which do not intuitively make sense as also transforming Z into something can be regarded as trivially transforming each such Z back into Z itself.

3 Quine on Fitting and Guiding

One of the most vexed and longstanding issues in the foundations of linguistics concerns the 'psychological reality' of grammars. Probably the most famous challenge to linguistics along these lines was posed by Quine (1972). In this section, I consider the Quinean challenge, and offer a measurement-theoretic response. Broadly speaking, I argue that rather than posing a deep problem for a psychological interpretation of grammars, Quine's observation helps to show what is right about such an interpretation.

Quine held that there was an important limit to what a grammar could reveal about human linguistic abilities. Here, briefly, is Quine's argument. For any given grammar, there will always be infinitely many other 'extensionally equivalent' grammars – i.e., grammars that produce the very same set of expressions as the original one. Thus, if these equivalent grammars can't be teased apart empirically, then none of them can be regarded as the 'correct' grammar. That is, no grammar can be held as 'guiding' a speaker's behavior, in Chomsky's sense of accurately reflecting the internal mental organization of a speaker's 'competence' (cf. below). Instead, a grammar can at most be said to 'fit' the speaker, in the sense that the speaker 'conforms' to it (Quine, 1972, 442).

Quine, it should be noted, saw the distinction between a grammar's guiding vs. merely fitting a speaker as a permanent and highly principled problem; cf. George, 1986. For him, the problem is not merely that at present the available data merely *underdetermine* which particular grammar should be accepted. Rather, even if the available data is a complete infinite set, there will still be infinitely many grammars consistent with it. Thus, since there will *always* be multiple grammars that accommodate the data, the question of which one is 'correct' is *indeterminate* – i.e., there is no correct answer.

In what follows, it will be preferable to consider a strong and updated version of Quine's argument. (Importantly, though, the basic argument below applies regardless of how one interprets Quine. The stronger version is merely a more interesting foil.) There are at least three changes to the historical argument to be made. First, Quine held that a rule *guides* a speaker only if the speaker 'knows the rule and can state it'; i.e., she '*observes* the rule' (Quine, 1972, 442). But the Chomksyan notion of guidance like the one characterized above is more relevant for present purposes. In particular, the above notion of guidance does not require that the speaker should be able to state the relevant rules, or have any other sort of conscious access to it – indeed, Chomsky and many others have frequently denied that any such access is quite implausible. (Quine, of course, would not tolerate this Chomskian midpoint (Quine, 1972, 444; cf. 446).[10] Fortunately, the main point to be defended does not depend on any particular such view, so for present purposes I adopt the stronger one.) Second, Quine's discussion focuses 'systems of English grammar'. But it is more common to consider a grammar to be any (normal) mature state of FHL, not just those that correspond to 'English'. Finally, Quine appears to regard two grammars as equivalent if they possess the same weak generative capacity – i.e., if they produce the same set of strings (of symbols, phonemes, etc.). He does not mention any other criteria (psychological, biological, theoretical, etc.) that are typically also considered relevant to the evaluation of a grammar's adequacy (e.g., Berwick et al., 2011, 4–11). These adjustments can still support a Quinean argument. For any grammar/theory of FHL, regardless of its conscious

---

[10] 'Chomsky would of course credit the native with a full and precise sense of grammaticality, this being of a piece with the native's purported fund of tacit rules – the native's purported bias even among extensionally equivalent grammars. Now this doctrine is interesting, certainly, if true; let me only mention again the crying need, at this point, for explicitness of criteria and awareness of method.' (Quine, 1972, 444, 446).

accessibility, there will always be infinitely many empirically equivalent distinct grammars of FHL. Thus, none of them can be regarded as more 'correct' than the others. In what follows, this is the argument to be considered.

The most well-known response to (either form of) Quine's argument is given in Chomsky, 1986, 248–257. There Chomsky argues that since linguistic theorizing is practiced in the same relevant ways as other scientific disciplines, we are thereby entitled to take seriously the explanatory posits of the resulting best theory. That is, the same rationale that legitimizes hypotheses about the molecular structure of benzene, the internal structure of a cell, the varying internal density of the earth, etc. similarly serves to legitimize hypotheses about the structure of human linguistic abilities. 'If our best theory accounts for Jones's behavior by invoking these rules and other elements, we conclude that they enter into Jones's behavior and guide it, that they play a 'causal role' in the sense of this discussion' (Chomsky, 1986, 249). Upon inspection, Chomsky argues, the issues involved in such attributions in linguistics 'are not crucially different from those that arise in any intellectual work of significance' (Chomsky, 1986, 252). This is a very common theme (Chomsky, 1976, 1995a, 2000, 2002, 2007, 2009, Chomsky and McGilvray, 2012). Indeed, one might think that taking seriously the empirical posits of a best theory is a big part of what it is to regard the theory as true. Thus, Chomsky maintains, if we are (tentatively) entitled to regard a theory of the human faculty of language as true, then we are (tentatively) entitled to regard the structure it posits as characterizing how a speaker's linguistic behavior is guided by that faculty.

Chomsky's reply tackles Quine's objection head-on: in ordinary scientific activity, the best theories are unproblematically regarded as describing the guidance – i.e., the theoretically relevant structure, often not directly observed – behind some empirical phenomenon. Since linguistics is similar to these other disciplines in the relevant ways, we are also justified in regarding its best theories as capturing the cognitive structure that in fact guides FHL. Thus, one extremely natural way of reading this reply is that we can eliminate (or at least reduce the importance of) the various equipotent grammars that each vie for the status of providing FHL's guiding structure. In short, the overarching goal of this reply involves finding a unique grammar with no Quinean empirical equivalents.

There is, however, an even better response to Quine. It begins by happily conceding that, as with nearly all other cases of measurement, there will be an infinitude of alternative empirically equivalent grammars, none of which can be favored over the others. In this sense, we grant Quine everything he wants. However, it doesn't follow that none of them guide the speaker in any useful sense. After all, *what they all have in common* may well supply many empirically rich and important details about the psychological characteristics of the speaker. That is, the meaningful structure of any particular grammar – what is invariant across all notational variants – may very well uncover much guidance, even though individual grammars contain additional empirically meaningless structure. Moreover, since what is common to notational variants is by hypothesis determined precisely by their empirical consequences, the potential for identifying those guiding features of a grammar is made as sharp as possible.

This last point is key. The relevant issue is *never* 'Does grammar G guide the speaker?' in the sense that G captures the true structure of FHL. Rather, the very best that can reasonably be hoped for is an answer to the subtler question '*Which features* of G guide the speaker?', which amounts to asking which features of G are empirically meaningful, insofar as G represents FHL.

An anonymous reviewer worries that empirically equivalent grammars might be 'too unlike to count as doing the same guiding'. But, by hypothesis again, notational variants differ only in their empirically meaningless properties, and agree on all their empirically meaningful ones. So the best shot any one them has at providing guidance is via their shared meaningful properties. After all, the latter, and only the latter, are determined by the empirical facts about FHL, at least as they are known, or hypothesized to be at the time; there is no replacement for basic empirical research. (In other words, if there's little shared structure amongst the set of empirically equivalent theories, then none of them have much empirical bite, and it's unclear in what sense any of them constitute an empirical theory at all.) Of course, different equivalent grammars may suggest different hypotheses for further exploration – but so might a television show about restaurants, a new statistical method, a discovery about some genetic process, etc. Some representations of temperature are more useful than others, and presumably the same holds for grammars. But this doesn't bear on the facts about what their empirical consequences are, or what features of a given grammar are empirically meaningful.

In short, in a sense that really matters, Quine's and Chomsky's views are consistent. Chomsky can embrace the fact that the linguist is 'knowingly and cheerfully up to his neck' in empirically equivalent grammars. But this is just an ordinary fact of scientific modeling. Rather than being a deep problem, it's an important aid in understanding the relevant properties of a model and its relation to the empirical phenomenon it models.

As a point of logic, the existence of empirically meaningless features is *equivalent* to the existence of multiple notational variants. So Quine's indeterminism as such amounts not to a deep philosophical problem, but to an injunction to understand a given representation's symmetry group, as specified by a uniqueness theorem like (2). This is clear in the case of temperature. Because of the meaningless features induced by the variability of the location and scale parameters, no one numerical temperature scale 'guides' temperatures, in the sense of capturing e.g., the empirically 'true' distance from the origin (there is no such truth). Instead, only *some* of the consequences of a favored representation (Fahrenheit, Celsius, Kelvin, etc.) are empirically meaningful. But this meaningful structure – namely, the affine ordering component – does capture important empirical facts about the structure of temperature. E.g., it predicts that temperatures are not periodic, with extreme heat being identical to extreme cold.

Alternatively, if a grammar is identified as its equivalence class of notational variants, then Quine's claim that there are many empirically equivalent theories is simply false: by hypothesis, there is exactly one. Moreover, in advance of characterizing the set of notational variants, it is similarly incorrect that a theory has been offered to be evaluated in the first place, much less the two he offers as rivals (Quine, 1972, 442).

In sum, rather than steadfastly defending all the structure and consequences of a particular grammar or 'best theory', we would do better to try to identify its meaningful empirical consequences. Indeed, this attitude is easily summed up in a general principle:

(IP) *Invariance Principle*: The interesting empirical question is *never* 'What does my particular favored formal grammar say about the matter?', but rather 'What does every notational variant of my favored formal grammar agree on about the matter?' That is, in linguistics as elsewhere, the interesting empirical content of a theory is that which is invariant across all notational variants that express it.

Although the pedigree of (IP) comes from mathematically tractable areas, its importance is the same for other areas like linguistics. Moreover, (IP) provides a simple and natural extension to the broader issue of the 'psychological reality' of grammar: the psychological reality that a correct grammar of FHL reveals is exactly that grammar's meaningful structure. Note, however, that talk of psychological reality still remains subtle and difficult. As generative linguists are wont to emphasize, their theories are not meant to capture how linguistic abilities are instantiated in the brain, or the means by which language is processed. Rather, the theorizing aims to capture a more abstract 'functional' level of the structure of linguistic competence (cf. Marr, 1982). Precisely what this level is (and how it relates to other levels of description) is not a simple matter. However, at this point, it helps illustrate how the comparison with temperature is especially apt. We saw above that only the bare affine structure of a numeric representation of temperature is meaningful. Although this affine numerical structure doesn't 'guide' temperatures in the sense of being causally responsible for the structure of the latter, it does capture a key empirical fact about temperature. In a similar sense, characterizing the structure of linguistic competence may reveal little about its neural or causal realization, but may nevertheless capture much important empirical information about the phenomenon at hand. Given the massive increase in complexity as we move from the model of temperature to FHL, it is natural to expect this latter body of information to be correspondingly rich and informative. (I am grateful to an anonymous referee for pressing me to clarify this point.)

Although Quine did not pose his objection using (IP) or anything like it, it seems that this is what lurks at its core. And in fact, Quine's general point remains, at least in principle, potentially devastating. This would be the case if the set of notational variants turned out to be so broad that it ruled out as meaningless virtually all the interesting consequences of a particular theory. (This, incidentally, is precisely the worry that, in a very different context, Anderson addresses regarding the possibilities for uncovering the nature of mental representations; Anderson, 1978, 278.) But the justification for such a claim requires substantive details about the particular case at hand, something which Quine never supplies and for which there is little evidence. What is called for here is, ideally, a uniqueness result corresponding to (2). In any case, these difficult matters of detail should replace sweeping Quinean pronouncements.

4 Biolinguistics: Simplicity and Behaviorism

This section considers how a measurement-theoretic perspective can illuminate some issues arising for the relatively new 'biolinguistic' program.

The currently popular trend of biolinguistics aims to explain how language might have evolved in human beings (Hauser et al., 2002, Chomsky, 2005, Boeckx, 2006, Fitch, 2010, Jackendoff, 2011, Di Sciullo and Boeckx, 2011, Berwick and Chomsky, 2011). Evolutionarily speaking, it is hard to explain the appearance of highly detailed, highly language-specific mental mechanisms. Conversely, it would be much easier to explain language's evolution in humans if it were composed of just a few very simple mechanisms. These considerations have increased attention to how much of language might be explicable by some very simple operations. The most prominent example of such an operation is Merge. For present purposes, Merge can be thought of as a computational rule that merely gathers two objects together, and labels the composite with one of them, this choice being determined by the morphological properties of the elements; e.g. Chomsky, 2005, 15.

(4) a. Merge(B, C) = {B, {B, C}} =X

b. Merge(A, X) = { X, {A, X }}, or = { A, {A, X }}

As (4) shows, iterated structures can take forms that we might choose to represent with the more familiar notation given in (3).

Thus, evolutionary considerations place a very high premium on the simplicity of the computational procedures that collectively compose a grammar. This contrasts with a more traditional approach that emphasizes descriptive adequacy, and thus licenses grammars composed of highly specific descriptive rules, such as the extreme example in (5); cf. Chomsky, 2005, 7, Berwick and Chomsky, 2011, 29 for other examples.

(5) Extraction from coordinate structures is impossible.

(Coordinate structures may be assumed to include, *inter alia*, constituents of the form [$_{XP}$ A *and* B].) On the one hand, if language operated by consulting rules like (5) in the construction of I-expressions, we would be able to easily and correctly predict the ungrammaticality (6a); cf. (6b).

(6) a. *Who did John and __ carry the piano upstairs together?

b. John and Mary carried the piano upstairs.

On the other hand, it is highly unclear how a principle like (5) might plausibly have evolved, unless it factors into some much simpler processes. Thus, by seeking formal grammars that are ultimately quite 'simple' in certain respects, it may be possible to obtain some form of cross-disciplinary theoretical economy. Moreover, as is frequently noted, by itself, the methodological goal of seeking out the simplest adequate theory often helps to identify the really crucial elements of any system; e.g., Chomsky, 2005, 10 , Berwick and Chomsky, 2011, 29.[11]

It should be noted that 'simplicity' is not specified antecedently. Rather, it is a proxy for some more complex characterization of what is biologically plausible, expedient, etc. As Berwick et al. put it, 'The idea is emphatically not that complex operations are biologically computed in surprisingly *efficient* ways. The hypothesis is rather that the core linguistic operations are simple enough to be computed by whatever biology underpins the generative processes that are exhibited by natural language grammars' Berwick et al., 2011, 13, emphasis added. That is, biologically plausible computational processes needn't be efficient in any abstract mathematical or intuitive sense. Nevertheless, it's a reasonable, though contingent, hypothesis that a general computational procedure like Merge in (4) is more likely to be biologically available (and explicable) than the much more specific and distinctively linguistic (5).

However assessed, a theory's simplicity is closely related to the measurement-theoretic issues under discussion. In fact, there are at least two reasons why biolinguists (and others) *require* a hypothesis about the uniqueness (cf. (2)) of a given theory. (Of course, assessing simplicity, like any other complex and challenging task, can always be done implicitly; but such inexplicitness is not normally a goal in the growth of scientific methods.[12]) After all, a representation's uniqueness determines its notational variants, which determines the meaningfulness of bits of the representation. But a complex bit of meaningless structure in some favored theory shouldn't count against the simplicity of the theory per se – especially when the relevant kind of simplicity is underwritten by biological considerations. Similar sentiments hold for the computational processes that operate on representations. Suppose e.g., a biologically simple process is characterized in some complex way, using much meaningless structure. Such a characterization might suggest the process is complex; but if the meaningless structure is identified as such, the ultimate assessment of the process may be quite different. Moreover, identifying extraneous, eliminable extra structure can often help to sharpen a theory, better exposing its true biological commitments. Thus, in this straightforward fashion, issues of

---

[11] Cf. also e.g. Forster and Sober, 1994, Forster, 2001, Baker, 2011 for general discussion of simplicity.

[12] As Dawes, Faust, and Meehl note, if the rule for calculating the total at the supermarket gets more complex than simply adding prices, we wouldn't expect informal human judgment to get better: 'Suppose instead that the supermarket pricing rule were, 'Whenever both beef and fresh vegetables are involved, multiply the logarithm of 0.78 of the meat price by the square root of

invariance can help identify the 'virtually conceptually necessary' operations (Chomsky, 2005, 10, Boeckx, 2006, chap. 3) that are required for it to operate as it does.

The second reason is this. Sometimes the most useful general representational format allows the *identification*, but not *elimination* of meaningless structure. This is obvious for temperature. Temperatures are typically represented with real numbers. But in so doing, some extra structure is inescapably used, namely the two parameters of an affine transformation (in other words, temperature is not a 'dimensionless quantity'). Moreover, (1) and (2) entail that there is no one unique correct representation, of which all others are affine transformations. Instead, *every* representation can perform this function (thus, they collectively form a 'regular' scale Roberts and Franke, 1976).

Like temperature, any stated grammar will almost surely *require* some meaningless structure in its representations (and this is the norm; cf. the passage above from Batchelder and Narens, 1977). Moreover, since language requires much more complex representations than ordinary quantitative phenomena, such meaningless structure may be summarily harder to control or even identify. Indeed, even in relatively simple cases, it is easy for practicing scientists to employ meaningless structure in erroneous inferences; cf. this footnote for a concrete example.[13] Thus, considerable measurement-theoretic caution should be used when characterizing or implementing a rule or other bit of structure into a grammar – even if the

---

twice the vegetable price'; would the clerk and customer eyeball that any better? Worse, almost certainly' Dawes et al., 1989, 1672.

[13] E.g, Sidowski and Anderson, 1967 claim to find a significant interaction between subjects' preference rankings of pairs consisting of (i) type of occupation (e.g., doctor, teacher, accountant), and (ii) type of city, as measured by its overall desirability. They offer an empirical (causal) interpretation of this interaction, writing that 'a teacher would tend to be in more direct contact with the socioeconomic milieu of a city', and thus more heavily influenced by it than those in other professions. But Krantz et al. showed that this interaction was, in the present terminology, empirically meaningless Krantz et al., 1971, 445–446. (More precisely, Sidowski and Anderson employed some interval-scale properties of their numeric representations of

former is 'virtually conceptually necessary'. Even the overall very best representations may unavoidably contain extra meaningless structure, and it may be hard to spot.

As the computational processes underlying language become simplified, it becomes increasingly plausible that their individual biological explanations are quite general, having little to do with language per se. That is, part of the evolution of FHL might have involved the recruitment of quite general processes whose own evolutionary explanations have little or nothing to do with language per se. In this sense, some properties of language may resemble the hexagonal shape of beehive cells. The hexagons result from compressing roughly circular walls together; thus, the explanation is ultimately geometric, and has little to do with bees as such. Chomsky calls such general explanations the 'third factor' of language design (Chomsky, 2005), supplementing the two more familiar influences of genetic endowment and (linguistically relevant) environmental influences. Obviously, such third factor explanations become more plausible the more that a computational procedure's language-specific features can be shown to be empirically meaningless.

Importantly, biolinguists aren't the only ones seeking these kinds of third-factor explanations: they also are an important component of some behaviorist programs. E.g., this sort of third factor explanation would be welcome to avowed behaviorists like William Uttal (Uttal, 2003, chap 3). In many respects, this is unsurprising: third-factor explanations involve the recruitment of very general, often individually rather simple, resources to explain seemingly complex mental phenomena. For the behaviorist, a process like Merge stands a much better chance of being identified with some basic types of physically characterized processes, which needn't themselves be regarded as mental. Thus, replacing complex, distinctively linguistic rules like (5) with more general ones like Merge would also advance the behaviorist's program. Of

---

merely ordinal empirical phenomena.) A remarkably small rescaling of the responses that course, it's far too early to suggest a substantive meeting of the minds between generative linguists and behaviorists. Nevertheless, it's worth briefly observing some methodological and practical similarities.

Even if physical, behavioristically acceptable, correlates of some putatively mental phenomenon can't be found, measurement-theoretic considerations of invariance may still be able to rule out a phenomenon as mental. For example, consider Zipf's law, which relates a word's rank-order in a body of text with its relative frequency. (Zipf's law states roughly that the nth most commonly occurring word in a body of text will occur with a frequency of about F/n, where F is the frequency of occurrence of the most common word; cf. Li, 1992.) This law was at one time interpreted as a distinctively psychological phenomenon, concerning how individuals comprehend relevant elements of a problem (Uttal, 2003, 51). But this distinctively psychological explanation collapsed when the same law relating ranks and frequencies started showing up all over the place. Zipf's law, it turns out, also correctly describes income change, volcanic behavior, animal length, randomly generated strings, and a host of other phenomena.

It is not currently known *why* the law holds when it does. However, the inductive support for some kind of third-factor explanation of the law is clear. Zipf's law holds as an invariant across an extremely diverse collection of physical circumstances. Thus, whatever (physical/mathematical/statistical) structure ultimately explains Zipf's law will almost surely have essentially nothing to do with psychological properties qua psychological properties. In other words, the diversity of nonpsychological examples shows that the distinctively psychological features that realize an instance of Zipf's law are themselves empirically meaningless in terms of the law.

---

preserved all preference rankings yielded a perfect additive structure, with no interactions.

(A reviewer objects that it 'may well be that psychological properties, or some of them, themselves have empirical properties in virtue of which Zipf's law holds.' In some sense this is trivially true: if a psychological mechanism realizes Zipf's law (or any other feature, for that matter), then *by definition*, it could only realize the law via its *ex hypothesi* psychological structure. But barring a coincidence on a monumental scale, the relevant *explanatory* features of that structure are shared by volcanoes, income change, etc. Thus, just as hexagonal cells have little to do with beelike properties as such, there is no serious reason to suppose that Zipf's law for word usage – or any other psychological phenomenon – has much to do with psychology per se. In short, there's a reason you won't be subscribing to *The Journal of Psycho-Vulcanology*.)

It often seems hard to imagine any sort of serious rapprochement between arch rationalists like Chomsky and behaviorism. But in terms of methods and goals, the two camps may be closer than the decades of firey rhetoric would suggest, even if their initial philosophical starting points are radically different. But, as measurement theorists might say, if methods and goals remain invariant across research programs, but philosophical motivations do not, so much the worse for the latter. This is not to suggest these two broad viewpoints must necessarily converge, or that they are at heart only one view. However, it is quite striking – and heartening – that even fiercest theoretical rivals at bottom share a broad (and perhaps broadening) foundation of principles of basic science.

5 Getting Started: Some First Steps

In this final section, I briefly address some practical aspects of implementing these measurement-theoretic considerations. Table 1 above lists a variety of (metaphysical) parallels between the measurement of temperature and language. However, there are several important (epistemological) differences. When it comes to language, the empirical data are only partially grasped, and no serious theory of FHL is even in the offing. Similarly, the formal representations of individual expressions are not just real numbers, and the notational variants between any (fragments of) theories of FHL are unlikely to be related in any simple way that can be characterized by a few varying numerical parameters. These differences are summed up in (7):

(7) In contrast to the simple case of temperature measurement:

1. the empirical phenomena regarding FHL are not axiomatized, but are only partially characterized, with new empirical information constantly appearing;
2. the grammatical representation of FHL is highly complex, and only partially understood;
3. the formal representations of individual expressions are not simple real numbers, but are formally much less tractable mathematical structures;
4. the notational variants of a grammar are collectively unknown, and most likely are related in complex and currently unknown ways.

The difficulties just listed are sobering. (7) entails that the measurement-theoretic aspects of FHL in Table 1 are presently not well-articulated enough to enable explicit formal characterizations of representation and invariance corresponding to (1) and (2) (This outcome mirrors that for similar issues in the philosophy of mind; Dresner, 2010, 419, 425, 435, Matthews, 1994, 131 Matthews, 2007, 118, 125, 126, 127, 195-6, Crane, 1990 Stalnaker, 1984, 9, Swoyer, 1991.) But it should be clear by now that this is a serious worry, since it means that those parts of a linguistic theory with empirical bite cannot presently be identified and distinguished from the theory's meaningless structure. (The latter is serious, of course, since it amounts to being unable to say just what the empirical content of the given empirical theory is.)

The sorts of difficulties scouted in (7) leave linguistic theorizing in a curious position. In particular, there are two very different answers to the general question of just what the content of a given grammar is. On the one hand, grammars are mathematically precise computational structures; in this sense they could hardly be clearer. On the other hand, it is quite unknown just what features of any given grammar are empirically meaningful; in this sense, matters are less obvious.

Despite these difficulties, headway is possible. One can begin by making the usual methodological simplifications and idealizations: attending only to manageable interesting 'bits' of a grammar, and similarly for empirical data, generalizations, etc.[14] (This is a major strength of much work in measurement theory; various bits of empirical phenomena can be very precisely carved out and studied in a piecemeal fashion. Trying to study all of the mind is a fool's errand, and probably trying to study, e.g., just the structure of human preferences is too. But real work can be done in trying to figure out why they sometimes appear not to be transitive.) Similarly, one can hazard inductive inferences to the effect that either some feature is empirically meaningful, or that a notational variant lacking it will be found (which is how the example of Zipf's law above operates).

It is also possible to force the issue and explore potential 'counterexamples' – i.e., relevant notational variants of bits of a grammar that lack some purportedly meaningful bit of structure. Perhaps the most well-known such example is one due to Ellis (1966). He showed that in the case of the extensive measurement of lengths (of straight rods, say), although combined lengths could be taken by abutting the lengths end to end, this was not necessary. Instead, the relevant axioms could also be satisfied by placing the rods at right angles, and taking the length of the resulting hypotenuse as their combined length; cf. Krantz et al., 1971, 87–88 for discussion.

There is no reason that all these strategies cannot be employed towards grammars. A natural first step might be to analyze some prominent hypothesized subcomponents of FHL and its representation. Merge, e.g., would be a natural place to start.

According to the characterization in (4), Merge is in one sense unordered; it simply forms sets. (Other processes, we may assume, determine the linear phonological structure Chomsky, 2005, 15.) But the two merged elements are not treated symmetrically, since one of them is projected as a 'label', determined by their respective morphological properties.[15] Thus far, then, *in advance of further empirical conditions*, in addition to (4), one might also represent the process as an ordered, noncommutative, nonassociative concatenation operation that always projects the first element, where the latter order is morphologically determined:

(8) a. $B \oplus C = B \oplus D = X$

b. $C \oplus B = Y$

c. $A \oplus X = A \oplus (B \oplus C)$

In (8), the unspecified morphological features specify the order in whch the two elements are input into $\oplus$ (e.g., (8b) might never occur, possibly for reasons unrelated to $\oplus$; if it does occur,

---

[14] Unsurprisingly, opinions differ drastically about what is manageable and what is tantamount to a 'theory of everything'; e.g., Chomsky, 2000, Jackendoff, 2002, Jackendoff, 2011.

[15] 'Each syntactic object generated contains information relevant to further computation. Optimally, that will be captured entirely in a single designated element, which should furthermore be identifiable with minimal search: its label, the element taken to be 'projected' in X-bar-theoretic systems. The label, which will invariably be a lexical item introduced by external Merge, should be the sole probe for operations internal to the syntactic object, and the only element visible for further computations.' Chomsky, 2005, 14

Y may be distinct from X). In (4), in contrast, such morphological features fully determine which element is projected. Structurally, (8) resembles a function that yields a weighted average of two numbers (cf. Luce et al., 1990, chap. 19 for extensive analysis of such representations). Needless to say, (8) appears rather different from (4); both, however, are much more plausible evolutionarily speaking than, say, (5). If (8) is to be ruled out as *empirically* inadequate for some reason, those reasons should be stated explicitly, ideally as axioms in a measurement-theoretic analysis.

For example, if, for whatever reason, the commutativity of the inputs (as in (4)) is felt to be empirically important, this should be explicitly stated: an appropriate binary function $*$ should satisfy $a * b = b * a$. (There will also be a number of other 'background' conditions to be stated, too, such as that $*$ is defined on some subset of pairs of the expressions, yielding as values only non-atomic (nonlexical) expressions; $*$ is 1-1, etc.). Obviously, (4) satisfies this commutativity condition, but (8) may not (it will iff X = Y). Similarly, to characterize the fact that only one element determines the projected 'label', we might proceed in two steps. First, we might define a relation R on expressions, so that a R b intuitively captures the idea that a determines the label in the concatenation (Merging) a*b. What empirical properties R possesses is an interesting question. Presumably it is reflexive and often asymmetric when $a \neq b$; being also transitive would be a boon to linguistics. With a little work, it could then form the basis of an equivalence relation $a \sim b$ that indicates that a and b project the same 'label'. With that in hand, the second step supplies an axiom:

(9) If $b \, R \, b'$, $c \, R \, c'$, and $b * b' \sim c * c'$, then $a * (b * b') \sim a * (c * c')$

The special case of (9) where b = c provides the needed invariance: the higher projection depends only on a and b, and is not affected by any b−dominated arguments. (If the 'same' label means merely equivalent (~), then (9) establishes something weaker than strict identity.) (4) satisfies the axioms as sketched so far. But some obvious questions for future work are what other functions besides (4) also satisfy (9), and what other features besides (9) are relevant (and what functions satisfy them).

On a methodological note, it's worth observing that this measurement-theoretic perspective can complement some more common forms of theorizing. E.g., rather than proposing some functional mechanism such as Merge, and subsequently exploring what roles it is able to serve, measurement theoretic considerations can be thought of as 'reverse-engineering' such mechanisms. That is, instead of simply proposing a complete mechanism, measurement-theoretic axioms can be thought of as putting constraints on what properties any such mechanisms must have. Measurement-theoretic axioms define only piecemeal features of mechanisms, rather than entire mechanisms; invariance results, which characterize the class of mechanisms that have these features, specify the available candidates. (And, of course, whenever this set contains unacceptable mechanisms, that means there is need for further (axiomatic) explicitness about what the needed mechanism(s) must be like.) Actually, Merge represents a curious midway point here. On the one hand, strictly speaking, no operation of Merge has ever been defined anywhere in generative linguistics. After all, the partial characterization in (4) does not determine which element is to be projected, and such a process may be quite complex. To the extent that Merge is used repeatedly in the generation of an I-expression, this resultant complexity can be enormous. On the other hand, Merge is typically stipulated (tacitly, perhaps) to be commutative, which in general is a highly nontrivial property of empirical operations. To my knowledge, this feature has never been defended empirically. (Moreover, attempted defenses based on intuitions of 'simplicity' are hard to make sense of in advance of understanding the role that the features of the inputs to Merge affect its subsequent outputs.)

This piecemeal approach can be of great service when implementing constraints that come nowhere near specifying any particular function. For instance, questions of the learnability of a grammar from appropriate background conditions might put certain weak constraints on the innate functions, or the nature of the functions that need to be partly specified during the acquisition process. Provided such constraints can be stated, they can act just like any other axioms to further constrain the empirically equivalent functions. Oftentimes such collections of constraints can yield fruitful empirical results even though no specific structure that satisfies all of them has been specified. (In the limiting case, sometimes it can be shown that collections of consistent constraints are collectively inconsistent, so that no structure can possibly satisfy them all.)

In short, measurement theory provides a means for explicitly working backwards, from needed features to the cognitive mechanism(s) that have them, rather than from proposed mechanisms to the features that they in fact have. Obviously, there is a great deal more work here to be done, in terms of filling out a characterization of Merge, studying its properties and its interactions with other features of FHL. I hope to take some initial steps in this direction in a later paper.

6 Conclusion

This paper has shown how, like scientific theorizing in general, linguistic theorizing admits of a measurement-theoretic treatment. Such an analysis distinguishes a grammar's meaningful structure concerning FHL from its meaningless structure that is an artifact of how the entire theory is articulated. It is important to theoretically distinguish these two kinds of structure, since a grammar's meaningless structure may not be eliminable from a statement of it. Moreover, there is no guarantee that the two kinds of structure can be easily distinguished; but failure to do so invites interpretations of empirical phenomena based on mathematical artifacts. Contra Quine, then, the existence of notational variants is not only unproblematic for understanding the guidance a given grammar supplies, it is utterly crucial. The key is that the alternative components of linguistic theories that also 'fit' the speaker supply the means by which the meaningful structure of a grammar is isolated. Thus, it is only through them that we can identify what guidance the grammar supplies. Moreover, as the final section shows, it is possible to take the kinds of small, explicit steps that better reveal just what one's favored theory really says.

Department of Logic and Philosophy of Science

University of California, Irvine

References

Anderson, J. 1978: Arguments concerning representations for mental imagery. *Psychological Review*, 85(4), 249–277.

Baker, A. 2011: Simplicity. In Zalta, E. N., editor, The Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/archives/sum2011/entries/simplicity/, summer 2011 edition.

Batchelder, W. and Narens, L. 1977: A critical examination of the analysis of dichotomous data. *Philosophy of Science*, 44(1), 113–135.

Berwick, R. C. and Chomsky, N. 2011: The biolinguistic program: The current state of its development. In Di Sciullo, A. M. and Boeckx, C., editors, *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*, 19–41. Oxford University Press, Oxford.

Berwick, R. C., Pietroski, P., Yankama, B., and Chomsky, N. 2011: Poverty of the stimulus revisited. *Cognitive Science*, 1–36.

Boeckx, C. 2006: *Linguistic Minimalism: Origins, Concepts, Methods, and Aims*. Oxford University Press, Oxford.

Chang, H. 2004: *Inventing Temperature: Measurement and Scientific Progress*. Oxford University Press, Oxford.

Chomsky, N. 1972: *Studies on Semantics in Generative Grammar*. Mouton, The Hague.

Chomsky, N. 1976: Problems and mysteries in the study of human language. In Kasher, A., editor, *Language in Focus: Foundations, Methods and Systems,* 281–357. D. Reidel Publishing Co., Dordrecht, Holland.

Chomsky, N. 1986: *Knowledge of Language*. Praeger, Westport, Conn.

Chomsky, N. 1995a: Language and nature. *Mind*, 104(413), 1 – 61.

Chomsky, N. 1995b: *The Minimalist Program*. MIT Press, Cambridge, MA.

Chomsky, N. 2000: *New Horizons in the Study of Language and Mind*. Cambridge University Press, Cambridge.

Chomsky, N. 2002: *On Nature and Language*. Cambridge University Press, Cambridge.

Chomsky, N. 2005: Three factors in language design. *Linguistic Inquiry*, 36(1), 1–22.

Chomsky, N. 2007: Of minds and language. *Biolinguistics*, 1, 9–27.

Chomsky, N. 2009: The mysteries of nature: How deeply hidden? *Journal of Philosophy*, 106(4), 167–200.

Chomsky, N. and McGilvray, J., editors 2012: *The Science of Language: Interviews with James McGilvray*. Cambridge University Press, Cambridge.

Churchland, P. 1979: *Scientific Realism and the Plasticity of Mind*. Cambridge University Press, Cambridge.

Collins, J. 2004: Faculty disputes. *Mind and Language*, 19(5), 503–533.

Crane, T. 1990: An alleged analogy between numbers and propositions. *Analysis*, 50, 224–230.

Davidson, D. 1989: What is present to the mind. In Brandl, J. and Gombocz, W., editors, *The Mind of Donald Davidson*, 3–18. Editions Rodopi, Amsterdam.

Dawes, R., Faust, D., and Meehl, P. M. 1989: Clinical versus actuarial judgment. *Science*, 243(4899), 1688–1674.

Di Sciullo, A. M. and Boeckx, C. 2011: Introduction: Contours of the biolinguistic research agenda. In Di Sciullo, A. M. and Boeckx, C., editors, *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*, 1–16. Oxford University Press, Oxford.

Dresner, E. 2004: Over-assignment of structure. *Journal of Philosophical Logic*, 33, 467–480.

Dresner, E. 2010: Language and the measure of mind. *Mind and Language*, 25(4), 418–439.

Ellis, B. 1966: *Basic Concepts of Measurement*. Cambridge University Press, London.

Field, H. 1981: Postscript. In Block, N., editor, *Readings in the Philosophy of Psychology*, 112–114. Harvard University Press, Cambridge, MA.

Fitch, W. T. 2010: *The Evolution of Language*. Cambridge University Press, Cambridge.

Forster, M. 2001: The new science of simplicity. In Zellner et al., 2001, 83–119.

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.

Fromkin, V., editor 2000: *Linguistics: An Introduction to Linguistic Theory*. Blackwell's, Oxford.

George, A. 1986: Whence and whither the debate between Quine and Chomsky? *The Journal of Philosophy*, 83(9), 489–99.

George, A. 1989: How not to become confused about linguistics. In George, A., editor, *Reflections on Chomsky*, 90–110. Blackwell's, Oxford.

Guttman, L. 1955: The determinacy of factor score matrices with implications for five other basic problems in common-factor theory. *The British Journal of Statistical Psychology*, VIIII, 65–81.

Hauser et al., 2002 Hauser, M., Chomsky, N., and Fitch, T. 2002: The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298, 1569–1579.

Jackendoff, R. 2002: *Foundations of Language*. MIT Press, Cambridge, MA.

Jackendoff, R. 2011: What is the human language faculty? two views. *Language*, 87(3), 586–624.

Krantz, D. H., Luce, R. D., Suppes, P., and Tversky, A. 1971: *Foundations of Measurement Volume 1: Additive and Polynomial Representations*. Academic Press, Inc., San Diego.

Laudan, L. and Leplin, J. 1991: Empirical equivalence and underdetermination. *The Journal of Philosophy*, 88(9), 449-472.

Li, W. 1992: Random texts exhibit Zipf's-law-like word frequency distribution. *IEEE Transactions on Information Theory*, 38(6) , 1842–1845.

Luce, R. D., Krantz, D. H., Suppes, P., and Tversky, A. 1990: *Foundations of Measurement Volume 3: Representation, Axiomatization, and Invariance*. Academic Press, Inc., San Diego.

Maraun, M. D. 1996: Metaphor taken as math: Indeterminacy in the factor analysis model (with discussion: *Multivariate Behavioral Researc*h, 31(4) , 517–538, 539–689.

Marr, D. 1982: *Vision*. Freeman, San Francisco.

Matthews, R. 1994: The measure of mind. *Mind*, 103(410) , 131–146.

Matthews, R. 2007: *The Measure of Mind: Propositional Attitudes and Their Attribution*. Oxford University Press, Oxford.

Moultin, J. and Robinson, G. 1981: *The Organization of Language*. Cambridge University Press, Cambridge.

Narens, L. 2002: *Theories of Meaningfulness*. Laurence Erlbaum Associates, Mahwah, NJ.

Narens, L. 2007: *Introduction to the Theories of Measurement and Meaningfulness and the Use of Symmetry in Science*. Laurence Erlbaum Associates, Mahwah, NJ.

Roberts, F. and Franke, C. H. 1976: On the theory of uniqueness in measurement. *Journal of Mathematical Psychology*, 14, 211–218.

Sidowski, J.B.,. and Anderson, N. 1967: Judgments of city-occupation combinations. *Psychonomic Science*, 7, 279–280.

Stalnaker, R. 1984: *Inquiry*. MIT Press, Cambridge, MA.

Suppes, P. 2002: *Representation and Invariance of Scientific Structures*. CSLI Publications, Stanford, CA.

Suppes, P., Krantz, D. H., Luce, R. D., and Tversky, A. 1989: *Foundations of Measurement Volume 2: Geometric, Threshold, and Probabilistic Representations*. Academic Press, Inc., San Diego.

Swoyer, C. 1991: Structural representation and surrogative reasoning. *Synthese*, 87(3), 449–508.

Uttal, W. R. 2003: *Psychomythics: Sources of Artifacts and Misconceptions in Scientific*

Psychology. Lawrence Erlbaum Associates, Mahwah, NJ.

Quine, W. 1972: Methodological reflections on current linguistic theory. In Davidson, D. and Harman, G., editors, *Semantics of Natural Language*, 442–54. D. Reidel, Dordrecht.

Zellner, A., Keuzenkamp, H. A., and McAleer, M., editors 2001: *Simplicity, Inference and Modelling*. Cambridge University Press, Cambridge.