DOI: 10.1111/phpe.12179

ORIGINAL ARTICLE

PHILOSOPHICAL PERSPECTIVES WILEY

Functionalism and tacit knowledge of grammar

David Balcarras

Department of Linguistics and Philosophy, Massachussetts Institute of Technology, Cambridge, Massachusetts, USA

Correspondence
David Balcarras, Massachussetts Institute of Technology, Cambridge, Massachusetts, USA.
Email: balc@mit.edu

Abstract

In this article, I argue that if tacit knowledge of grammar is analyzable in functional-computational terms, then it cannot ground linguistic meaning, structure, or sound. If to know or cognize a grammar is to be in a certain computational state playing a certain functional role, there can be no unique grammar cognized. Satisfying the functional conditions for cognizing a grammar G entails satisfying those for cognizing many grammars disagreeing with G about expressions' semantic, phonetic, and syntactic values. This threatens the Chomskyan view that expressions have such values for speakers because they cognize grammars assigning them those values. For if this is true, semantics, syntax, and phonology must be indeterminate, thanks to the indeterminacy of grammar-cognizing (qua functional-computational state). So, the fact that a speaker cognizes a grammar cannot explain the determinate character of their language.

1 | INTRODUCTION

You understand this sentence. And this one, assigning it structure, meaning, and sound. You interpret both as having syntactic, semantic, and phonological properties, as you do this one and as you will for others to follow. You interpret them effortlessly, thanks in part to your linguistic competence: your tacit knowledge of grammar. For you are wired to interpret language according to the grammar you tacitly know. You can parse and understand this entire essay, and even speak it for yourself as you are doing now, presumably in silence in inner speech. These feats of linguistic performance are made possible by your competence. Moreover, your interpretations of the sentences of this essay, even though they may depart from the author's, are guaranteed to be correct at least in your idiolect. For the syntactic, semantic, and phonological properties that expressions have in your idiolect are precisely the ones assigned to them by the grammar you tacitly know; they have those properties for you *because* they are assigned by your grammar. As such, the character of your idiolect is fundamentally psychological, fixed by a state of your mind. Likewise for all idiolects. And thus language itself, being a motley sum of idiolects, has a mental core. All of this Chomsky taught us.[1]

But is this story true? The question matters because, if its answer is 'Yes', this has significant philosophical consequences. One is that semantic determinacy is possible where we have long been told indeterminacy is inevitable. For if there is a fact about which grammar I tacitly know or (better) 'cognize' (Chomsky 1975b, p. 165), and if this fixes the meanings of words in my mouth, then it is settled what those words mean. And so it is settled whether an expression in my idiolect means the same thing as, or is synonymous with, an expression in someone else's. Also, given that my cognized grammar works compositionally, the semantics of my whole language is fixed in one stroke, including the meanings of expressions hitherto thought to be difficult or impossible to pin down, such as the meanings of subsentential expressions, unused and unusable expressions, and sentences about the unknowable. And a final consequence is that linguistics must answer to what is in the mind. For a semantic, syntactic, or phonological theory about a speaker's idiolect is accurate only insofar as it agrees with the grammar they cognize. So meaning is in the head after all.

But none of this is foregone. As I argue in this paper, whether the Chomskyan story is true depends crucially on what it is to cognize a grammar. More specifically, if grammar-cognizing is to be understood in broadly functional-computational terms, then language's meaning, structure, and sound cannot be psychologically grounded in it. To show this, I set forth a standard version of functionalism as applied to grammar-cognizing, on which to cognize a grammar is to be in a certain computational state playing a certain causal-functional role. But once the details of this view are on the table, it looks like there can be no such thing as *the* grammar a speaker cognizes. Because if a speaker's mind/brain satisfies the functional conditions for cognizing a grammar G, it must also satisfy the conditions for cognizing many other grammars, many of which disagree with G about the syntactic, semantic, and phonological properties of expressions. Thus, the fact that a speaker cognizes a grammar cannot explain the determinate character of their idiolect or the correctness of the interpretations they assign to expressions.

There are two interpretations of this result. The first is that at least part of what Chomsky taught us is false; grammar-cognizing is not up to the task of fixing language's meaning, structure, and sound. And the second is that grammar-cognizing is nothing like the functional-computational account says it is; the cognizing relation cannot be functionally analyzed and is instead *sui generis* or analyzable in some other way.

I prefer the first interpretation. But the aim of this paper is not to establish that it is the correct one. I do not know which is correct. Nevertheless, it is an advance if we can show that one of the two must be. And so that is what I intend to show.

2 | WORK FOR A THEORY OF COGNIZING

For the Chomskyan, which grammar a speaker cognizes determines the correct syntactic, semantic, and phonological interpretations of the expressions of their idiolect or language. So let us get clearer on what it would take for a *theory* of grammar-cognizing (hereafter just 'cognizing') to accommodate this. The claim that cognizing does this determining work can be abstractly summarized as the claim that which grammar a speaker cognizes fixes which language they know or have. More precisely, the claim is this:

> GROUND: For any speaker *s* and language *L*, it is necessary that *s* has *L* iff *s* cognizes some grammar *G* for *L*.

Some clarifications are in order. First, in order for GROUND to have any plausibility at all, '*s*' and '*L*' had better not range over all possible speakers and languages. Chomskyans are only concerned with human speakers and humanly acquirable natural languages.[2] The possibility of a Martian speaking a language without cognizing a grammar, or of a human knowing a programming language like C++ without cognizing a grammar for it, are supposed to be compatible with GROUND. Similarly, if a human can use, read, or write a second language without cognizing a grammar for it, GROUND is not supposed to be refuted. And it is also not supposed to be refuted by infant speakers or speakers with severe linguistic impairments who might not cognize grammars. Thus, by 'a speaker', we had better mean something like a normal adult human native speaker, and by 'a language' we had better mean a natural language. And, finally, I leave open whether 'necessary' in GROUND should be read as nomological or metaphysical necessity. But, to make potential counterexamples to GROUND harder to come by, I will read it nomologically.

Next let us draw out some consequences of GROUND. Following Lewis (1975), let us model a language *L* as a function from expressions to meanings, and treat the following biconditional as true by definition: An expression *e* has the meaning *m* for a speaker *s* iff *s* has some language *L* such that $L(e) = m$. With this in place, the following metasemantic principle follows from GROUND:

> METASEMANTICS: Necessarily, *e* means *m* for *s* iff *s* cognizes some *G* for some *L* such that $L(e) = m$.

Moreover, if we individuate expressions in terms of their form—if we model any *e* as a pair of a phonological (or phonetic) form and a logical form, ⟨PF, LF⟩, following Chomsky (1995b)—then GROUND also entails metasyntactic and metaphonological principles:

> METASYNTAX: Necessarily, *e* has the logical form *LF* for *s* iff *s* cognizes some *G* for some *L* such that *e* is an expression of *L* and *LF* is the second member of *e*.[3]

> METAPHONOLOGY: Necessarily, *e* has the phonological form *PF* for *x* iff *s* cognizes some *G* for some *L* such that *e* is an expression of *L* and *PF* is the first member of *e*.[4]

I like this combination of principles. And I recommend that any analysis of language-having, or of the 'actual-language relation' (Schiffer 1993), be put to more than metasemantic work in this way.[5] So I take GROUND to entail METASEMANTICS, METASYNTAX, and METAPHONOLOGY.

Clearly then, if GROUND is true, we can give (at least nomologically) necessary and sufficient conditions for a core set of linguistic facts in terms of cognizing. And if cognizing can be analyzed in non-semantic/-syntactic/-phonological terms, then GROUND will take us some way towards an illuminating theory of meaning and metaphysics of language. But whether this is so depends on what cognizing is. So let us turn to what cognizing might be.

3 | SCHEMATIC THEORIES OF COGNIZING

What is it to cognize a grammar? Many philosophers of linguistics have offered roughly the same incomplete answer: to cognize a grammar is to be in a certain brain state, such that which grammar one cognizes is determined in some unspecific way by unspecified neural goings-on. For Hornstein (2009) cognizing has an "implementation in brain-like material" and "must ultimately be embedded in brain circuitry" (p. 3); cognizing is "realized" in "neural circuity" and "supervene[s]" upon "brain structure" or "the computational circuitry and wiring that the brain embodies" (pp. 156, 9). For Ludlow (2011) there are "lower level physical processes" and a "low level biophysical state upon which [cognizing] supervenes" (pp. xvii, 63).[6] For Pietroski (2018) cognizing is "biochemically realized" by "biologically implemented generative procedures" (pp. 69, 8). For Rey (2020) cognizing is a "computational system" somehow "*realized* or *implemented*" in "some local physical properties in the brain" (pp. 23, 114, 368).

And we find similar schematic appeals to the neural basis of cognizing all throughout Chomsky: cognizing is "physically *realized* in a finite human brain" (Chomsky and Halle 1968, p. 6), "must *structurally correspond to* some features of brain mechanism" (Chomsky and Katz 1974, p. 364), has a "neural *basis*" (Chomsky 1975b, pp. 8, 40), is "*coded*" and "*realized* in some arrangement of physical mechanisms" (1982b, p. 15), must somehow "*arise* in the mind/brain" 1988, p. 3), is "*determined* by the nature of the mind/brain" (p. 36), and has a neural "*implementation*" (2016, p. 110); we also read that "one task of the brain sciences is to determine what it is about [the] brain *by virtue of which*" we cognize grammars, and that facts about the brain "*explain*" and are "*responsible for*" cognizing, and that the "mind/brain" somehow "*yields*" cognizing (1986, pp. 22, 39, xxvi) (emphases mine).[7]

But none of these authors explicate which specific realization or implementation relations they have in mind. In fact, it is difficult to find any explicit account of the specific realization relation in which cognizing is supposed to stand to the brain.[8] This is striking, given the huge amount of attention that realization or dependence relations have been given in the philosophy of mind more broadly.

But perhaps there is a reason for this. For Chomsky argues that "the relation of brain and mind", and so of brain and cognizing, "is a problem of the natural sciences", and so, presumably, not a matter for philosophical speculation (1986, p. 40). He motivates this by imagining a case in which we discover a "complex of neural mechanisms" that "corresponds" to a grammar G with principles $P_1, \ldots, P_n$ but not to G' with "equivalent" principles $Q_1, \cdots, Q_n$, allowing us to infer that those mechanisms give rise to cognizing G not G' (p. 40). But this only shows that it is an empirical question *which* cognizing states depend on *which* neural states. The nature of this dependence relation remains up for philosophical investigation. For how is it that possessing a certain neural mechanism is metaphysically sufficient for cognizing a grammar with certain principles? This is a philosophical question.

So the task is to make sense of how cognizing depends on the brain, with an eye to assessing whether GROUND is true or plausible. To do so, one might naturally reach for a *functionalist* account of cognizing. This would make the most sense of schematic talk of cognizing being 'realized'. But talk of 'implementation', on the other hand, suggests that cognizing is a computational state, a state of implementing a grammar (or a program or set of functions to which it corresponds). So is cognizing supposed to be functional or computational?

I suggest a view on which it is both. In what follows, I present a standard version of functionalism as applied to cognizing, on which to cognize a grammar is to stand in a neural relation to that grammar that plays a certain functional role. But this view contains a gap: it tells us nothing about how it is possible for a speaker to stand in a neural relation to a grammar. I suggest that this gap is best filled by a computational account of how speakers' brains implement grammars. The combined computational-functional view then says: to cognize a grammar is to stand to a grammar in a neural computational implementation relation that plays the cognizing-role. I then argue that this view cannot be squared with GROUND, survey some potential solutions to this problem, and conclude that these solutions are unsuccessful.

4 | A COMPUTATIONAL-FUNCTIONAL THEORY OF COGNIZING

Functionalism about cognizing is, roughly, the thesis that to cognize a grammar is to be in some state that plays a certain functional role, or stands in certain causal relations to inputs to the mind, its outputs, and to other mental and neural states. But let us build up to a more precise formulation.

4.1 | The functional realization of cognizing

Let us grant that cognizing is a psychological relation posited by some true and complete Chomskyan psycholinguistic theory T, such that many sentences of T contain 'cognizes'. And let us grant that T is not too radically different from the best such theory on offer today, at least in how it characterizes cognizing. And suppose that T characterizes cognizing functionally, specifying the characteristic relations of causal or counterfactual dependence in which states of cognizing stand.[9] Suppose also that T is a theory only of *speakers*, those normal adult human native speakers of natural languages in the Chomskyans' purview.

If we conjoin T into a single sentence, it might look something like this:

> For any speaker s and grammar G, if s were to cognize G and …, then the state of s's cognizing G would, *ceteris paribus*, causally enable events of s perceiving sentences generated by G to cause events of s interpreting those sentence as having the syntactic, semantic, and phonological properties assigned to them by G, and ……… .

with '………' filled in with many more specifications of cognizing's definitive causal powers, and '…' filled in with a specification of the background conditions under which cognizing typically manifests those powers.[10] We might speculate, following Chomsky (1992), that what replaces '…' will say that a state of cognizing is "embedded" or "integrated into performance systems" for "articulation" and "interpretation" (p. 213).[11]

Also note the '*ceteris paribus*' qualifier. Because T is a psychological 'special science' theory, the law-like generalizations it contains had better tolerate exceptions. This is especially important because T is Chomskyan theory of linguistic competence rather than performance.[12] For given speakers' performance limitations, T would be false if '*ceteris paribus*' were removed. As Lewis (1994) would put it, as functional theory like T should characterize cognizing's "typical causal role" (p. 301).[13]

Now, if we let 'C' be a binary predicate constant abbreviating 'cognizes', we can represent the whole open-sentence following 'For any speaker s and grammar G' above as '$\mathsf{T}[\mathsf{C}, s, G]$'. And then we can represent T as follows:

(1) $\forall s \forall G.\mathsf{T}[\mathsf{C}, s, G]$

Henceforth, 'T' refers to the interpreted sentence (1). And let us say that a relation R *realizes* T, or plays the cognizing-role, just if, for some s and some G, $\mathsf{T}[R, s, G]$. And we can also say that R realizes T in a speaker s just if, for some G, $\mathsf{T}[R, s, G]$.[14]

With all of this on the table, we can state a standard formulation of functionalism, as applied to cognizing, as follows:

> FUNCTIONALISM:
>
> Necessarily, x cognizes G iff there is a unique relation R such that $\mathsf{T}[R, x, G]$ and $R(x, G)$.

Why accept FUNCTIONALISM? A classic argument is metasemantic.[15] It is argued that a theoretical term like 'cognize' has its meaning fixed by being implicitly defined by a theory like T, more or less as envisioned by Ramsey (1929), Carnap (1961), and Lewis (1970). To say that T implicitly defines 'cognize' is to say at least that what 'cognizes' actually means is the same as what it would have meant if it had been introduced by explicitly defining it as follows: Let 'x cognizes G' mean the same as 'x stands to G in the R such that $\mathsf{T}[R, x, G]$'! If this is what 'x cognizes G' means, then, given a Russellian analysis of definite descriptions, FUNCTIONALISM follows.[16]

Now, we can interpret the Chomskyan hypothesis that cognizing is neurally realized in functionalist terms as follows:

> HYPOTHESIS:
>
> For each speaker s, some neural relation N uniquely realizes T in s.

Given FUNCTIONALISM, this means that, for each speaker s, there is a neural relation such that s is in a state of standing in that relation to whatever grammar they cognize, and that state stands in all the causal relations that T says states of cognizing stand in. But we should take care to distinguish three relations in which a speaker can stand to a grammar:

(P1) $\lambda x \lambda G.\, N(x, G)$
(P2) $\lambda x \lambda G.\, \mathsf{T}(N, x, G)$ and $N(x, G)$
(P3) $\lambda x \lambda G.\, \exists! R.\, \mathsf{T}[R, x, G]$ and $R(x, G)$

Given FUNCTIONALISM, (P3) is logically equivalent to $\lambda x \lambda G.x$ cognizes G. And (P2) and (P1) are what Shoemaker (1981) would call the *total realization* and *core realization* of an instance of cognizing, respectively.

We can now see how FUNCTIONALISM allows us to explain how cognizing can be neurally realized. Given that HYPOTHESIS is true, for each speaker s, there is a (P1)-type neural relation N such that their standing in it to G is sufficient for their standing in a (P2)-type relation to G, which is in turn sufficient for their standing in (P3) to G, which, by FUNCTIONALISM, logically entails that s cognizes G.

But there is something missing. We have explained how a speaker can cognize G by standing in a neural relation to G. But how is *that* possible?[17] The core realization of cognizing is just as mysterious as cognizing itself. What are these neural relations supposed to be?

4.2 | Implementing a grammar

The most natural way to answer this question, and to complete the functional explanation of how cognizing is neurally realized, is to think of the core realization of cognizing in computational terms. That is, I suggest that the Chomskyan hypothesis that cognizing is neurally realized should be elaborated as follows:

> COMPUTATIONAL HYPOTHESIS:
>
> For each speaker s, some neural *computational implementation* relation N uniquely realizes T in s.

What motivates this is that theoretical talk of 'cognizing' is widely interpreted as being made at Marr's computational level of description.[18] Cognizing is said to involve computing or implementing a set of functions (or program or other abstract whatnot) corresponding to a grammar. If so, the theory that speakers cognize grammars is like Marr's (1982) theory that the human visual system detects edges by computing the Laplacian of a two-dimensional Gaussian distribution of the retinal input (pp. 54–61).

Here is a more precise statement of COMPUTATIONAL HYPOTHESIS:

(CH) For each speaker s, there is some neural relation N such that (i) N uniquely realizes T in s, and, (ii) necessarily, $N(s, G)$ just if some system in s's brain implements G.[19]

FUNCTIONALISM plus COMPUTATIONAL HYPOTHESIS make up a computational-functional theory of cognizing. Of course, to round it out, we need to ultimately supplement it with an account of what it is for a brain mechanism to implement a grammar.[20] But, insofar as we are confident that the brain *is* an implementing organ, and that implementation is explicable in neural terms, we can be confident that the computational-functional theory of cognizing is viable.

I think FUNCTIONALISM plus COMPUTATIONAL HYPOTHESIS is an attractive pair of views that many Chomskyans explicitly or implicitly endorse or should endorse. Unfortunately, as I will argue next, their combination cannot accommodate GROUND.

5 | A PROBLEM FOR FUNCTIONALISM ABOUT COGNIZING

But before turning to that argument, I want to address the worry that whether FUNCTIONALISM can accommodate GROUND is not pressing because any functionalist thesis is bound to be false or at least overwhelmingly objectionable. And as the literature does abound with objections to theses like FUNCTIONALISM, I should say why I do not take myself to be beating a dead horse.

First, many objections to functionalist theses only target so-called 'analytic', 'commmonsense', or 'a priori' functionalism. These do not apply to FUNCTIONALISM because the theory T is not an a priori or analytic folk theory. FUNCTIONALISM is rather a version of 'scientific' or 'a posteriori' functionalism, dubbed 'psychofunctionalism' by Block (1978).[21] And, second, although there is no shortage of objections to this or that version of psychofunctionalism, it not at all clear how many of them apply to FUNCTIONALISM. Most (if not all) objections to psychofunctionalism are of one of two types: (a) objections that psychofunctionalism mischaracterizes conscious, personal-level, or folk mental goings-on, such as belief, desire, or perception; or (b) objections that psychofunctionalism about a certain mental property is out of line with our everyday judgments or ordinary intuitions about its actual or possible extension. But type-(a) objections do not apply to FUNCTIONALISM because cognizing is a subpersonal, non-conscious mental state. And type-(b) objections do not apply because 'cognizing' is a *pure* theoretical term unused by the folk. And so there is no reason to expect that our ordinary judgments or intuitions about cognizing are accurate.[22]

For these reasons, FUNCTIONALISM is hardly dead out of the gate. Rather, it is a natural regimentation of how cognizing is schematically characterized by those who posit it. And it is a direct consequence of a plausible and influential account of the metasemantics of theoretical terms like 'cognize'.[23]

So it is significant if true that FUNCTIONALISM cannot accommodate GROUND.

5.1 | The problem of too many realizers

My main argument that FUNCTIONALISM is in tension with GROUND stems from an initial worry that it may be very difficult, if not impossible, for T to be uniquely realized by a neural relation in a speaker. The worry is *not* that T might be multiply realized by distinct neural relations in a single speaker relating them to the same grammar; this worry is familiar and easily remedied by simply dropping 'unique' from FUNCTIONALISM and COMPUTATIONAL HYPOTHESIS. Rather, it is an argument that, if some N realizes T within a speaker s and $N(s, G)$, then there is bound to be a multitude of distinct neural relations $N_1, N_2, \ldots$ relating s to other grammars $G_1, G_2, \ldots$ all of which play the cognizing-role in s. If so, then COMPUTATIONAL HYPOTHESIS is false, and FUNCTIONALISM entails that no one cognizes a grammar. But we cannot drop 'unique' to solve *this* problem. For once it is dropped, FUNCTIONALISM massively overgenerates. It entails that each s cognizing some G for their language L also cognizes a multitude of other grammars, many of which wildly disagree with G about the syntactic, phonological, and semantic properties of expressions of L. Given GROUND, this means that expressions of a speaker's language will have a multitude of wildly varying *but equally correct* interpretations. Semantic indeterminacy thus returns, but this time with a vengeance, bringing both syntactic and phonological indeterminacy in its wake. This is unacceptable. Even a Chomskyan who might have learned to live with a little semantic indeterminacy would not stomach phonological indeterminacy. So, FUNCTIONALISM fails to accommodate GROUND.

Now for the details.[24] Suppose that there is some neural relation N realizing T in some s. So, there is a grammar G such that $\mathsf{T}[N, s, G]$. And say this is witnessed by $\mathsf{T}(\mathbf{N}, \mathbf{s}, \mathbf{G})$ (where '$\mathbf{N}$', '$\mathbf{s}$', and '$\mathbf{G}$' are constants). But now consider the relation $\mathbf{N}^*$ defined as follows:

$$\mathbf{N}^*(x, G) \text{ iff } \mathbf{N}(x, f(G)).$$

where f is some bijection from the set of grammars onto itself such that $f(G) = G'$ only if G and G' generate the same sentences and assign them the same meanings. f maps a grammar to what I will call a 'sententially equivalent' grammar.[25] Now, as I will show, it could easily turn out that $\mathsf{T}(\mathbf{N}, \mathbf{s}, \mathbf{G})$ entails $\mathsf{T}(\mathbf{N}^*, \mathbf{s}, f^{-1}(\mathbf{G}))$, which would in turn entail that $\mathbf{N}^*$ realizes T in s. But in that case, because $\mathbf{N} \neq \mathbf{N}^*$, T is not uniquely realized in $\mathbf{s}$ by $\mathbf{N}$. Thus, generalizing, in no possible world is T uniquely realized in a speaker.

To see how $\mathsf{T}(\mathbf{N}, \mathbf{s}, \mathbf{G})$ might entail $\mathsf{T}(\mathbf{N}^*, \mathbf{s}, f^{-1}(\mathbf{G}))$, or why $\mathbf{N}^*$ might play the cognizing-role in **s** if **N** does, let us simplify massively and pretend T is such that a triple satisfies $\mathsf{T}[R, s, G]$ just if they satisfy (2).[26]

(2) If it were the case that $R(s, G)$ and …, then, *ceteris paribus*:
 (i) for any sentence S and meaning m, if G generates S and assigns m to S, then: if s were to perceive S, then s would interpret S as meaning m.

Now, $\mathsf{T}(\mathbf{N}, \mathbf{s}, \mathbf{G})$ is equivalent to the following closure or instance of (2):

(3) If it were the case that $\mathbf{N}(\mathbf{s}, \mathbf{G})$ and …, then, *ceteris paribus*:
 (i) for any S and m, if **G** generates S and assigns m to S, then: if **s** were to perceive S, then **s** would interpret S as meaning m.

But because $\mathbf{N}(\mathbf{s}, \mathbf{G})$ is definitionally equivalent to $\mathbf{N}^*(\mathbf{s}, f^{-1}(\mathbf{G}))$, (3) entails (4):[27]

(4) If it were the case that $\mathbf{N}^*(\mathbf{s}, f^{-1}(\mathbf{G}))$ and …, then, *ceteris paribus*:
 (i) for any S and m, if **G** generates S and assigns m to S, then: if **s** were to perceive S, then **s** would interpret S as meaning m.

And because **G** and $f^{-1}(\mathbf{G})$ are sententially equivalent, (4) is equivalent to (5):

(5) If it were the case that $\mathbf{N}^*(\mathbf{s}, f^{-1}(\mathbf{G}))$ and …, then, *ceteris paribus*:
 (i) for any S and m, if $f^{-1}(\mathbf{G})$ generates S and assigns m to S, then: if **s** were to perceive S, then **s** would interpret S as meaning m.

But (5) is just an instance of (2). And so (5) entails that $\mathsf{T}(\mathbf{N}^*, \mathbf{s}, f^{-1}(\mathbf{G}))$, or that $\mathbf{N}^*$ plays the cognizing-role in **s**. Thus, if **N** realizes T in **s**, then so must $\mathbf{N}^*$. And the same can be said for each relation N^*_h defined such that $N^*_h(x, G)$ iff $\mathbf{N}(x, h(G))$ for each bijection h from a grammar to a sententially equivalent grammar. All of them must realize T in **s** if **N** does. If this is true, this is a problem.

Now, of course, T in its full complexity will not be such that $\mathsf{T}[R, s, G]$ looks like (2). But the problem arises even if $\mathsf{T}[R, s, G]$ looks more like this:

(2*) If it were the case that $R(s, G)$ and …, then, *ceteris paribus*:
 (i) for any sentence S and meaning m, if G generates S and assigns S meaning m, then: if s were to perceive S, then s would interpret S as meaning m, and
 (ii) for any sentence S and logical form LF, if G generates S and $LF \in S$, then: if s were to perceive S, then s would parse S as having the logical form LF, and
 (iii) for any sentence S and phonological form PF, if G generates S and $PF \in S$, then: if s were to perceive S, then s would hear S as having the phonological form PF, and ……

That is, the problem arises so long as replacing 'G' with '$f^{-1}(G)$' in (i), (ii), (iii), and so on makes no difference to its satisfaction. And for (i)–(iii) this is so.

So, the question becomes: does $\mathsf{T}[R, s, G]$ (and so T) contain a component (n) such that, when it has a true closure $\mathsf{T}(\mathbf{N}, \mathbf{s}, \mathbf{G})$, it would be rendered false if we replaced '$\mathbf{G}$' with '$f^{-1}(\mathbf{G})$' in its (n) part? Here is a candidate:

(x) for any subsentential expression e and meaning m, if G generates e and assigns m to e, then: if s were to perceive e, then s would interpret e as meaning m

Now, because $f^{-1}(\mathbf{G})$ may very well disagree with $\mathbf{G}$ in its assignment of meanings to subsentential expressions (and we can suppose that it does), swapping '$\mathbf{G}$' for '$f^{-1}(\mathbf{G})$' in (x) in $\mathsf{T}(\mathbf{N}, \mathbf{s}, \mathbf{G})$ would make it false. The modified sentence says that $\mathbf{s}$ would typically interpret $f^{-1}(\mathbf{G})$-expressions as having their $f^{-1}(\mathbf{G})$-meanings if it were that $\mathsf{T}(\mathbf{N}, \mathbf{s}, \mathbf{G})$. But that is wrong, given that $\mathsf{T}(\mathbf{N}, \mathbf{s}, \mathbf{G})$ truly says that $\mathbf{s}$ would in that case typically interpret those expressions as having their distinct $\mathbf{G}$-meanings instead.

But there are two problems with this fix. First, it is hard to see how T could be true if it included (x). Do we typically grasp the meanings of arbitrary *sub*sentential expressions generated by the grammar we cognize? Elsewhere, I have argued that we do not (Balcarras 2023, pp. 10–12). Consider that if someone said to me the noun phrase 'snake next to you' on its own, I would not *ceteris paribus* interpret it having a subsentential semantic value. I would interpret it as elliptical for '*There is a* snake next to you' and as having the semantic value of that sentence. But I do not need to lean on this argument here.

The second and worse problem for the (x)-fix is that, even if T includes (x), we can modify our definition of $\mathbf{N}$'s permuted cousin $\mathbf{N}^*$ and show again that if $\mathbf{N}$ realizes T in $\mathbf{s}$ then so does $\mathbf{N}^*$. Again, suppose $\mathbf{N}$'s realizing T in $\mathbf{s}$ is witnessed by $\mathsf{T}(\mathbf{N}, \mathbf{s}, \mathbf{G})$. But this time define $\mathbf{N}^*$ as follows: $\mathbf{N}^*(x, G)$ iff $\mathbf{N}(x, i(G))$, where i is a bijection such that $i(G) = G'$ only if (a) G and G' are sententially equivalent, *and* (b) for any subsentential expression e less than 100,000 words long (or with less than 100,000 constituents), G and G' both generate e and assign e the same meaning. If we pretend that an expression's being 100,000 or more words long is the point at which it becomes practically impossible for a speaker to use it, we can say that i maps grammars to *practically equivalent* grammars.

Now, because practical equivalence entails sentential equivalence, $\mathbf{G}$ and $i^{-1}(\mathbf{G})$ are sententially equivalent. So we can run the argument that $\mathsf{T}(\mathbf{N}, \mathbf{s}, \mathbf{G})$ entails $\mathsf{T}(\mathbf{N}^*, \mathbf{s}, i^{-1}(\mathbf{G}))$ just as before, up until we face the reply that $\mathsf{T}[R, s, G]$ contains (x). Because $\mathsf{T}(\mathbf{N}, \mathbf{s}, \mathbf{G})$, which now looks like (6), we get (7) by definition:

(6) If it were the case that $\mathbf{N}(\mathbf{s}, \mathbf{G})$ and …, then, *ceteris paribus*: ……
(x) for any e and m, if $\mathbf{G}$ generates e and assigns m to e, then: if $\mathbf{s}$ were to perceive e, then $\mathbf{s}$ would interpret e as meaning m ……

(7) If it were the case that $\mathbf{N}^*(\mathbf{s}, i^{-1}(\mathbf{G}))$ and …, then, *ceteris paribus*: ……
(x) for any e and m, if $\mathbf{G}$ generates e and assigns e meaning m, then: if $\mathbf{s}$ were to perceive e, then $\mathbf{s}$ would interpret e as meaning m ……

The crucial question is whether (8) is also true:

(8) If it were the case that $\mathbf{N}^*(\mathbf{s}, i^{-1}(\mathbf{G}))$ and …, then, *ceteris paribus*: ……
(x) for any e and m, if $i^{-1}(\mathbf{G})$ generates e and assigns m to e, then: if $\mathbf{s}$ were to perceive e, then $\mathbf{s}$ would interpret e as meaning m ……

I say it is. To see why, it is best to explain what is wrong with a good-sounding yet bad argument that (8) is false. The bad argument is that, given (6), there will be counterexamples to (8). Let e be a subsentential expression with 100,001 constituents, a long description formed by stacking nouns: 'the swan boat race trophy case lock key polish bottle lid…'. Given (6), the thought is that if **s** were **N**-related to **G** (in a way integrated with their performance systems), then **s** would interpret e as having its **G**-meaning, not its i^{-1}(**G**)-meaning. So, (8) is false.

Unfortunately, if this style of argument refutes (8), then it also refutes (6). (6) also has counterexamples. If e' is a noun 300 trillion words long generated by **G**, **s**'s perceiving e' would not cause s to understand e' as having its **G**-meaning, for **s** could not completely parse e'. And, less fancifully, **s** might regularly misinterpret expressions of their language as not having their **G**-meanings. But this is not inconsistent with (6) because it is *ceteris paribus* generalization. Likewise, (8) is not refuted by pointing out that **s** would not interpret e or e' as having its i^{-1}(**G**)-meaning. And so it is not clear whether there are counterexamples of the right kind to refute (8), counterexamples that would refute (8) when read as a non-strict special science law. (And if somehow there are, we can just redefine i using '300 trillion' instead of '100,000', and presumably eliminate them.) So, again, **N*** seems destined to play the cognizing-role in **s** if **N** does.

Generalizing, if $N(s, G)$ and $\mathsf{T}[N, s, G]$, then $N_i^*(s, i^{-1}(G))$ and $\mathsf{T}[N_i^*, s, i^{-1}(G)]$, for *each* N_i^* defined such that $N_i^*(x, G)$ iff $N(x, i(G))$, one relation for each bijection i from a grammar to a practically equivalent grammar. Thus, T cannot be uniquely realized in a speaker, and a speaker cannot stand to a grammar in a relation uniquely realizing T in them. And so Computational Hypothesis is false, and Functionalism entails that no one cognizes a grammar.[28] As I said above, we can dodge this by dropping 'unique' from these theses and endorsing:

> Functionalism*:
>
> Necessarily, x cognizes G iff there is some relation R such that $\mathsf{T}[R, x, G]$ and $R(x, G)$.
>
> Computational Hypothesis*:
>
> For each speaker s, some neural computational implementation relation N realizes T in s.

But this makes things worse. Given the argument above, Functionalism* entails that if $N(s, G)$ and $\mathsf{T}[N, s, G]$, then s cognizes G and also $i^{-1}(G)$ for each i. So if I cognize G and it assigns all the correct semantic, syntactic, and phonological properties to the expressions of my language, I must also cognize each $i^{-1}(G)$. But many of these will be "bent grammars" (Lewis 1992b), assigning wild incorrect semantic interpretations to subsentential expressions 100,000 words or longer. One might assign the description 'the swan boat race trophy case lock key polish bottle lid…' $= e$ the same meaning it assigns to 'the cat'. Thus, given Ground, it follows that e and 'cat' are synonymous for me. But this is incredible.

And it gets worse. For suppose instead we take a neural relation N realizing T and define a bent relation N_d^* such that $N_d^*(x, G)$ iff $N(x, d(G))$ for each bijection d meeting this condition: $d(G) = G'$ only if, for any expression $\langle \mathrm{PF}, \mathrm{LF} \rangle$ with less than 100,000 constituents, G generates $\langle \mathrm{PF}, \mathrm{LF} \rangle$ just if G' does. Then, a parallel argument shows that if I cognize G, I must also cognize each $d^{-1}(G)$. But many of these will be syntactically and phonologically bent, generating bent $\langle \mathrm{PF}, \mathrm{LF} \rangle$ pairs. One might generate $e'' = \langle \mathrm{PF}', \mathrm{LF}' \rangle$, where PF' is the phonological form of $e =$ 'the swan boat race trophy case lock key polish bottle lid…' and LF' is the logical form of 'and'. Thus, given

GROUND, it follows that e'' is an expression of my idiolect, a connective that sounds like e. But that is not so. So we cannot combine FUNCTIONALISM* and GROUND.

But perhaps there is a way out of this mess. Let me now turn to some potential solutions.

6 | POTENTIAL SOLUTIONS?

6.1 | The 'competence' solution

The Chomskyan might argue that T makes some generalizations about what cognizing causes that are exceptionless, not *ceteris paribus*. Perhaps the *ceteris paribus* parts of T specify the causal relations in which cognizing typically stands to various 'performance events' (understandings, parsings, hearings), while T also says that cognizing *always* causally yields what I will call 'states of competence'. That is, perhaps cognizing G only *ceteris paribus* (diachronically) causes a speaker to understand G-expressions as having their G-interpretations, but it *always* (synchronically) causes a speaker to tacitly know that G-expressions have their G-interpretations, independently of whether it is integrated or interfacing with their performance systems.

In more detail, the proposal is that T will include and so entail a semantic generalization like (9) (along with similar syntactic and phonological generalizations):

(9) For any s and G, if s were to cognize G, then, for any e and m, if G generates e and assigns m to e, then s would know that e means m.

For if T contains (9), then because it is to be read as exceptionless, we cannot replace 'G' with '$i^{-1}(G)$' in $\mathsf{T}[R, s, G]$ whenever it is satisfied and expect it to remain satisfied. And so the cognizing-role specified by T will be strong enough to not be played by any permuted neural relation in which cognizers of G stand to $i^{-1}(G)$.

To assess this solution, note first that the semantic 'knowledge' that (9) says cognizing yields is, like cognizing itself, a theoretical posit of T. Thus 'know' in (9) should be replaced with something less familiar. Replace 'know' with the coinage 'postcognize', to remind us that postcognizing is posterior to cognizing. Like 'cognizing', 'postcognizing' is plausibly implicitly defined by T. And so what it is to postcognize something is captured by a functionalist thesis like FUNCTIONALISM. Indeed, perhaps cognizing and postcognizing should be functionally interdefined, just as belief and desire are typically interdefined by functionalists about folk psychology.[29]

But if this is right, the problem rearises. For whenever a pair of non-bent relations jointly play the cognizing/postcognizing-role, a multitude of pairs of bent relations will also jointly play it. And so nothing is solved by speculating that T contains (9).

To see this, suppose the (9)-part of $\mathsf{T}[R, s, G]$ is instanced by the following, where $\mathbf{N}$ plays the cognizing-role and $\mathbf{R}$ plays the postcognizing-role in $\mathbf{s}$:

(10) If it were that $\mathbf{N}(\mathbf{s}, \mathbf{G})$, then, for any e and m, if $\mathbf{G}$ generates e and assigns m to e, then it would be that $\mathbf{R}(\mathbf{s}, [e \text{ means } m])$.

Here '[e means m]' refers to what I will neutrally call an 'object of competence', that which (9) says we postcognize whenever we cognize a grammar assigning m to e, be it a proposition, sentence, or whatever. Now, to get the problem going again, we need to single out a relation $\mathbf{R}^*$ such that, given (10), (11) follows:

(11) If it were that $\mathbf{N}^*(\mathbf{s}, i^{-1}(\mathbf{G}))$, then, for any e and m, if $i^{-1}(\mathbf{G})$ generates e and assigns m to e, then it would be that $\mathbf{R}^*(\mathbf{s}, [e \text{ means } m])$.

To this end, let v_i map $i^{-1}(\mathbf{G})$-expressions to their $i^{-1}(\mathbf{G})$-meanings and let v map $\mathbf{G}$-expressions to their $\mathbf{G}$-meanings. Next, take a bijection k from the set of objects of $i^{-1}(\mathbf{G})$-competence (the set including $[e \text{ means } m]$ iff $v_i(e) = m$) to the set of objects of $\mathbf{G}$-competence such that:

$$k([e \text{ means } v_i(e)]) = [e \text{ means } v(e)]$$

And then define $\mathbf{R}^*$ as follows:

$$\mathbf{R}^*(x, [e \text{ means } m]) \text{ iff } \mathbf{R}(x, k([e \text{ means } m]))$$

We can show that, when $\mathbf{R}^*$ is so defined, (10) entails (11). This is because, given (10), if $\mathbf{N}^*(\mathbf{s}, i^{-1}(\mathbf{G}))$, (12) follows:

(12) If $i^{-1}(\mathbf{G})$ generates e and assigns m to e, then $\mathbf{R}^*(\mathbf{s}, [e \text{ means } m])$.

For suppose otherwise. That is, suppose that, for some e' and m', $v_i(e') = m'$ and $\neg\mathbf{R}^*(\mathbf{s}, [e' \text{ means } m'])$. By definition, it follows that $\neg\mathbf{R}(\mathbf{s}, k([e' \text{ means } m']))$. But $k([e' \text{ means } m']) = [e' \text{ means } v(e')]$. So $\neg\mathbf{R}(\mathbf{s}, [e' \text{ means } v(e')])$. But that is inconsistent with (10). Thus, given (10), (12) must be true if $\mathbf{N}^*(\mathbf{s}, i^{-1}(\mathbf{G}))$. So, given (10), (11) follows.

So, $\mathbf{N}^*$ and $\mathbf{R}^*$ must jointly play the cognizing/postcognizing-role if $\mathbf{N}$ and $\mathbf{R}$ do. Thus, it is no solution to our problem to speculate that T contains something like (9). For even if we do, bent relations will still realize T in any speaker in which straight relations realize T. The problem recurs.

In light of this, the Chomskyan might change tack and say that T actually employs the ordinary word 'know' in (9), not a theoretical term like 'postcognize'. Perhaps T contains parts like (9) that say that cognizing yields genuine propositional knowledge about the correct semantic, syntactic, and phonological interpretations of all expressions of speakers' languages. If so, perhaps the problem does not immediately rearise.

But there are two issues with this. First, it is implausible that speakers know the contents of all of the theorems of the grammar they cognize.[30] Indeed, Chomsky aims to accommodate this in coining 'cognize' to replace 'know' in characterizing 'knowledge' of language.[31] But note that even if speakers *do* happen to possess vast semantic knowledge as a matter of contingent empirical fact, that is neither here nor there. What matters is whether cognizing is (causally) sufficient for such knowledge. For if it is not, then playing the cognizing-role cannot require yielding this knowledge, and so T cannot include (9). But cognizing is not *on its own* sufficient for semantic knowledge. Semantic knowledge requires semantic belief, and our semantic beliefs must be realized in systems external to whatever mechanisms in our brains compute the grammars we cognize. The 'belief system' will be one of the performance systems with which cognizing interfaces. And so (9) should really look more like (9*) if 'know' means ordinary knowledge.

(9*) For any s and G, if s were to cognize G and …, then, *ceteris paribus*, for any e and m, if G generates e and assigns m to e, then s would know that e means m.

with '…' specifying that s's cognizing G is appropriately hooked up with s's belief-forming mechanisms, and with '*ceteris paribus*' reinserted (for those mechanisms might not operate in an atypical situation). But, of course, adding (9*) to T is no solution to our problem.

The second issue with thinking that 'know' means ordinary knowledge in (9) is that, even if (9) stays as-is, our problem plausibly rearises anyways, but in a different way. If a speaker cognizing G must causally thereby know and so believe that e means m whenever G assigns m to e, then these beliefs cannot be *explicit* beliefs; they cannot be each realized by a token mental representation, such as a token of Mentalese in the speaker's belief-box. For there are infinitely many semantic theorems of G, but a finite speaker's head can only contain finitely many Mentalese tokens. So, the semantic beliefs yielded by cognizing must be somehow implicit. But if so, the most natural explanation of how cognizing yields these beliefs, and so of how it yields vast semantic knowledge, will look something like this: When a speaker cognizes G, this yields finitely many representation-based beliefs in G's finitely many semantic axioms.[32] And because those axioms logically or deductively entail G's semantic theorems, it follows that the speaker implicitly believes and knows, or is in a position to know, those semantic theorems. For they are entailed by and so implicitly 'contained in' the content of their total belief-state.

Suppose this is what explains how $N(s, G)$ suffices for s's knowledge of G's theorems when N realizes T in s. The problem is that we can now specify a new problematic bijection j such that $N_j^*(s, j^{-1}(G))$ (where N_j^* is defined as above) suffices for s's knowledge of $j^{-1}(G)$'s theorems in the same way. Let j now be a bijection such that $j(G) = G'$ only if G and G' are intensional equivalents: the conjunction of G-axioms is intensionally equivalent to the conjunction of G'-axioms. Now, if s knows G's axioms and so is positioned to know G's theorems, then s is thereby also positioned to know $j^{-1}(G)$'s axioms (because they are equivalent to G's axioms) and so is also positioned to know $j^{-1}(G)$'s theorems. Thus it looks like whenever the (9)-part of $\mathsf{T}[R, s, G]$ is satisfied, it will remain satisfied when its unbound instances of 'G' are replaced with '$j^{-1}(G)$' (and the rest of $\mathsf{T}[R, s, G]$ will remain satisfied because intensional equivalents are practical equivalents). But we cannot allow that each $j^{-1}(G)$ is cognized whenever G is. Any realistic candidate for a grammar I cognize has infinitely many intensional equivalents that I do not cognize, such as grammars stuffed with billions of inert axioms expressing necessary truths. Thus I think nothing is solved by adding (9) to T.

6.2 | The computational solution

An alternative solution is to say that the neural relations realizing T in speakers do not *just so happen* to be computational implementation relations. Perhaps the Chomskyan should instead say that cognizing is necessarily computational, building it into T that cognizing a grammar requires implementing it. For then they can argue that if some N realizes T in s, it does not follow that every N_i^* also does. They can argue this because, if T says cognizing is computational, then, while $N(s, G)$ and $\mathsf{T}[N, s, G]$ still entail each $N_i^*(s, i^{-1}(G))$—and now entail that s implements G—they may not entail that s implements every $i^{-1}(G)$. And so they may not entail that N_i^* realizes T in s. This entailment will not go through, if, for instance, what it is for a system to implement a grammar is for its physical causal-structural organization to be isomorphic to the formal or derivational structure of that grammar. This is because G and $i^{-1}(G)$ will or may have different derivational structures. And so implementing G may not suffice for implementing $i^{-1}(G)$.

I have three reasons to not go in for this solution. First, I intentionally kept the hypothesis that cognizing is a computational relation separate from T. I did so because we should not assume that the cognizing-role cannot possibly be played by a non-computational neural relation. For it could turn out that the grammars we cognize are not computable.[33] So this solution requires too strong a stance on the content of T.

Second, because the solution appeals to *causal structuralism* about implementation, this raises a new overgeneration problem.[34] For any grammar, there is a multitude of distinct grammars with the same derivational structure. Consider, for any G, the grammar G^* formed by swapping out some atomic elements from G's lexicon for others. The swap might preserve derivational structure. And, given their distinct lexicons, G and G^* generate distinct expressions and determine distinct languages. Moreover, if G has a rich lexicon, there are many ways to lexically modify G into a distinct grammar with the same derivational structure.[35] So, if causal structuralism about grammar-implementation is true, then, if a mechanism in a speaker's brain implements a grammar, it thereby also implements many other grammars for languages other than theirs. This is an excrescence.

For Chalmers (1994), this consequence of causal structuralism is a feature not a bug. He admits "there is no canonical mapping from a physical object to "the" computation it is performing [...] To this very limited extent, the notion of implementation is 'interest-relative' " (p. 397).[36] But the Chomskyan will not want to say that which grammar I compute is interest-relative. And they should not admit that we compute potentially infinitely many distinct grammars. The reason is not that this is excessive. Rather, the issue is that the excess undermines the computational explanation of the core, neural realization of cognizing. Recall that the functional account of cognizing requires that we make sense of how a speaker can stand in a neural relation to a *particular* grammar. We cannot make sense of this by saying that they causal-structurally implement *it*. That would be misleading. In truth, given causal structuralism, they implement a perhaps infinite class of grammars and non-grammars—programs, automata, and so on—all with isomorphic structures.

Now, the Chomskyan might reply by denying causal structuralism, and claim instead that causal-structural isomorphism is necessary but not sufficient for implementation. But I suspect this may only postpone the problem until a different theory of implementation is selected.[37]

Be that as it may, here is the third and strongest reason to not endorse the computational solution. Although the derivational structure of a grammar G will surely differ from *many* of its practical equivalents, it will not differ from enough of them to hold off disaster. For we can show that *some* grammars do have practical equivalents with the same derivational structure that nonetheless 'go wacky' beyond the 100,000-word limit. To see this, imagine a possible speaker Floyd whose language L^* is bent and who cognizes its bent grammar G^*. L^* is practically equivalent to English, except in L^* every definite description 100,000-words or longer means the same as 'the cat'. So 'the swan boat race trophy case lock key polish bottle lid...' means the same as 'the cat' for Floyd. And we might also imagine that Floyd's brain contains a system whose causal organization is isomorphic to G^*.[38]

Now, because Floyd's language is already bent, this means it will have bent practical equivalents with the same derivational structure. For consider the grammar G^{**}, practically equivalent to G^*, but which interprets definite descriptions 100,000-words or longer as meaning the same as 'the mouse'. But because G^* and G^{**} have exactly the same derivational shape, Floyd's brain will satisfy the necessary causal-structural conditions for implementing both of them.[39] And so nothing will prevent Floyd from cognizing G^{**} given that he cognizes G^*.

Thus, even if T is modified to include the claim that speakers implement grammars, and even if implementing requires causal-structural isomorphism, this is not enough to block my argument that we cannot combine FUNCTIONALISM* and GROUND. For, given FUNCTIONALISM*, there are possible speakers who cognize and implement a grammar for their language and thereby must also cognize and implement grammars for languages other than their own, contra GROUND. And, for all we know, we may be speakers in a similar situation.

6.3 | The 'mirror constraint' solution

Building on the computational solution, another strategy is to argue that cognizing has stronger causal-structural requirements that ensure that a cognizer of a grammar *never* satisfies the requirements for cognizing its bent equivalents. This argument could be made based on the account of tacit knowledge of grammar developed by Evans (1981), Davies (1987), and Peacocke (1989).[40] On their view, for a speaker to cognize G is for their mind/brain's causal structure to 'mirror' G's structure, where this requires that they are in mental states corresponding to the axioms and rules of that grammar with causal powers to transition them, in the right conditions, into mental states corresponding to the theorems of that grammar. If this view is right, the Chomskyan can argue that T entails that cognizing requires mirroring; and that even if a speaker like Floyd cognizes a bent grammar G^* for their bent language L^*, they will mirror G^* but not its practical equivalent G^{**}, and so any bent neural relation in which Floyd stands to G^{**} will not realize T in them.

To assess this, let us say G^*'s bent rule is (R1) but G^{**}'s bent rule is (R2):

(R1) If e is a definite description with 100,000 or more constituents, then e means *the cat*.
(R2) If e is a definite description with 100,000 or more constituents, then e means *the mouse*.

Because Floyd mirrors G^*, Peacocke would argue that "some mechanism or algorithm in [Floyd] *uses the transition-type* expressed by the rule [(R1)]", which means that Floyd "is in a state with this causal power: that of causing [Floyd], if [he] is in a suitable [state] with [an] informational [content] of the form ['e is a definite description with 100,000 or more constituents'], to move into a state with an informational content of the form ['e means *the cat*']" 1989, p. 116).[41] If that is right, then the Chomskyan can argue that no mechanism in Floyd uses the transition-type expressed by (R2), and so Floyd does not mirror or cognize G^{**}.

But this solution fails for reasons similar to those that frustrated the competence solution. To see this, we can think of the generalized solution as the claim that T includes (14):[42]

(14) For any s and G, if s were to cognize G, then, if '$\phi \rightarrow \psi$' is a rule of G, then, if s were to be in a suitable state with the content ϕ, then s would be in a suitable state with the content ψ.

But what does 'a suitable state' mean in (14)? What, for Peacocke, is meant by a 'suitable state with an informational content'?

Here a dilemma arises. If what is meant is something like *postcognizing*, a theoretical attitude towards objects of competence posited and implicitly defined by T, then adding (14) to T will fair no better than adding (9) to T, which, as I argued above, solves nothing. On the other hand, if 'suitable states' are beliefs, then (14) is false. Here is a counterexample to (14) read as a thesis about belief: Suppose I take a phonology class and am taught some phonological axioms that just so happen to be the phonological axioms of the grammar I cognize. I now believe these axioms. But even though I cognize a grammar with rules for deriving phonological theorems from those axioms, I would not be caused to believe all of those theorems. So (14) cannot be a thesis about belief.

I read Peacocke as endorsing the first horn of my dilemma: the suitable states are functionally specified by T in terms of their causal connections to cognizing and other states. He claims that a suitable state with, for example, the content specifying the meaning of 'man' must "be connected with [its] subject's possession of the concept MAN, with the ground of his ability to recognize the word 'man', and with his general grasp of predication" (p. 113). Clearly, a belief that 'man' means *man* (or whatever) might not meet these requirements. Unlike beliefs, suitable states must be "connected" with a speaker's performance systems, with states underlying their capacities for understanding and thought (concept possession) and linguistic perception (word recognition). Indeed, Peacocke suggests that suitable states "function" to "allow the subject to participate, now as perceiver, now as producer," playing a causal role "in the perception and the production of sentences" (p. 118). And while he says that it "is a fascinating and difficult question what the nature of the required connections is," it must be that "no state which can exist without [these required connections] can be identified with" a suitable state (p. 115).

All this strongly suggests that to be a suitable state is to play a certain functional role. And that, like the postcognizing-role, the suitable-state-role is presumably intertwined with the cognizing-role. But if so, then whenever (14) is instanced by straight relations **N** and **S** as follows,

(15) If $\mathbf{N}(\mathbf{s}, \mathbf{G})$, then, if '$\phi \to \psi$' is a rule of **G**, then, if it were that $\mathbf{S}(\mathbf{s}, \phi)$, then it would be that $\mathbf{S}(\mathbf{s}, \psi)$.

we will no doubt be able to find bent relations $\mathbf{N}^*$ and $\mathbf{S}^*$ that we could swap in, with $i^{-1}(\mathbf{G})$, to get a bent instance of (14). So the 'mirror constraint' solution does not help.[43]

6.4 | The 'naturalness' solution

The final solution I will consider is to invoke the notion of 'naturalness', 'eligibility', or 'joint-carvingness', taking a cue from Lewis (1983, 1984) and Sider (2011).[44] The idea is to claim that the permuted, bent neural relations speakers stand in to bent grammars are somehow too 'gerrymandered' and 'grueish' to suffice for language-making instances of cognizing.

I think there are two ways to make good on this. The first is to reformulate GROUND using naturalness such that cognizing a bent grammar is insufficient for having the language of that grammar. And the second is to use naturalness to reformulate FUNCTIONALISM so that a neural relation relating a speaker to a bent grammar does not suffice for cognizing even if it plays the cognizing-role in that speaker. I will consider each strategy in turn and argue that neither is promising.

6.4.1 | Adding naturalness into the Chomskyan metasemantics

Suppose we concede that cognizing a grammar entails cognizing its practical equivalents. To stave off disaster, we might then say that among all the grammars a speaker cognizes, only *one* fixes the correct interpretations of expressions in their mouth, or, equivalently, only one fixes which language is theirs: the 'best' or most natural or 'straightest' grammar.[45] This amounts to modifying Ground as follows:

> Ground$^+$: Necessarily, s has L iff, for some grammar G, s cognizes G, G is a grammar for L, and G is the most natural grammar s cognizes.

I see three problems with Ground$^+$.

The first problem is a challenge: how exactly should we evaluate a grammar's naturalness? A related question has been explored in other contexts, in assessing Lewis's view that the actual grammar of the language of a population, insofar as there is one, is the "best" grammar that conforms to their conventional use of it.[46] But most of this discussion only addresses how to evaluate the bestness or naturalness of a grammar's semantic component. For instance, perhaps the best semantics for a language L is, roughly, the referential semantics that assigns L's expressions the most joint-carving worldly referents. But even if this is right, it tells us nothing about how to evaluate a grammar's syntactic and phonological components. But these need to be evaluable for naturalness too if Ground$^+$ is to help.

Lewis (1969) proposes that grammars be evaluated by various theoretical virtues. Straight grammars are "simple, natural, reasonable, easy, good", while bent grammars are "complicated, artificial, contrived, difficult, bad" (p. 198). But combining this proposal with Ground$^+$ is quite anti-Chomskyan. For Chomsky, which language a speaker has is supposed to be fixed by their psychology. But if Ground$^+$ is true and grammars are better or worse according to their theoretical virtuousness, then which language a speaker has is not fixed by their psychology. It is fixed by a mixture of psychology plus theoretical considerations of an arguably pragmatic, aesthetic, or otherwise interest-relative nature, with the latter doing all the 'fixing' work in singling out which grammar among the multitude cognized by the speaker gets the privilege of fixing their language.[47]

And even if the Chomskyan holds out for an alternative account of a grammar's bestness, they will face the second problem with Ground$^+$. This problem is that anyone who likes Ground$^+$ should like a similar principle about how the rules of a speaker's language are fixed by the best grammar they cognize. But this principle is false.

In more detail, if the Chomskyan thinks that the most natural grammar a speaker cognizes fixes their language, then they should also think that this grammar fixes other linguistic facts about that speaker, such as facts specifying their language's grammatical rules. Because a grammar determines a unique set of rules while a language does not, if cognizing *only* pins down a speaker's language, no unique set of the rules will be pinned down as *the* rules of their language. But the Chomskyan should avoid the result that a speaker's language has a multitude of distinct, conflicting rules as much as they should avoid the result that a speaker has a multitude of conflicting languages. For this reason, if they go for Ground$^+$, they should go for Rules$^+$:

> Rules$^+$: Necessarily, the rules of s's language are $R_1, \ldots, R_n$ iff, for some grammar G, s cognizes G, the rules of G are $R_1, \ldots, R_n$, and G is the most natural grammar s cognizes.

But Rules$^+$ is inconsistent with the empirical possibility that the rules of a speaker's language are the ones drawn from a grammar low in the naturalness ordering of grammars they cognize.

To illustrate, suppose s cognizes a minimalist grammar G (Chomsky 1995b, Collins and Stabler 2016), and so also cognizes its practical equivalents. Among these will be a very different kind of grammar, a *multiply context-free grammar* G'. This is because each minimalist grammar is 'strongly equivalent' to a multiply context-free grammar.[48] Strong equivalence entails practical equivalence. So G and G' are practically equivalent and generate the same language. But they have different rules. As Stabler (2013) shows, G''s rules will be "exponentially larger", and this additional complexity makes G' more gruesome (p. 2, Appendix A). So perhaps G is more natural than G'.

But even if so, it is possible that, although s cognizes both G and G', G''s rules are the actual rules of s's language. Following the argument from Chomsky (1986) discussed above, we could empirically discover neural mechanisms in s's head corresponding to G''s rules, enabling us to infer that G''s rules have pride of place for s. Such a discovery is nomologically, not just epistemically, possible. And the Chomskyan should be open to its possibility. So they should not endorse RULES$^+$. And so they should not endorse GROUND$^+$.

Finally, the third problem is that GROUND$^+$ is incompatible with the possibility of a speaker having a bent language L^* determined by a bent grammar G^* they cognize, like the speaker Floyd discussed above. For, given GROUND$^+$, if s has L^*, then L^* must be the language of the most natural grammar s cognizes. But if s cognizes G^*, they must also cognize its more natural practical equivalents, and thereby have the language L determined by the most natural practical equivalent of G^*, G. So L^* cannot be the language of the most natural grammar s cognizes. And so s cannot have L^*.

But it *is* possible for a speaker to cognize a bent grammar and thereby have a bent language. There could be a speaker such that the psycholinguistic and neural evidence confirms the hypothesis that G^* is the grammar of their language. Again, this evidence might point to neural correlates of the bent rules or axioms of G^* that are responsible for its going haywire for expressions longer than 100,000 words (even though these neural correlates never 'kick in'). So we might reasonably hypothesize that L^* is their language. And so we should not rule out this possibility by advancing GROUND$^+$.[49]

6.4.2 | Adding naturalness to the functional analysis of cognizing

Another naturalness-based solution begins by modifying FUNCTIONALISM* as follows:

> FUNCTIONALISM$^+$:
>
> Necessarily, x cognizes G iff there is some relation R such that $R(x, G)$ and $\mathsf{T}[R, x, G]$ and R is the most natural relation realizing T in s.[50]

The full solution is then to claim that whenever a speaker cognizes a grammar G that is *the* correct grammar of their language, they will stand in some neural relation N that realizes T in them that will always be more natural than any permuted neural relation N_i^* defined as above. If so, then even if each N_i^* realizes T in that speaker, by FUNCTIONALISM$^+$, they will not thereby *cognize* each incorrect bent grammar $i^{-1}(G)$ to which they are N_i^*-related.

To motivate this solution, one would need to establish the following principle:

(16) Necessarily, if $\mathsf{T}[R, s, G]$ and $R(s, G)$ and G assigns an *incorrect* interpretation to an expression of s's language, then R is not the most natural relation realizing T in s.

For if (16) is false, then it is possible for the most natural R realizing T in a speaker to relate that them to an *incorrect* grammar. And so, given FUNCTIONALISM^{+}, it is possible for a speaker to cognize an incorrect grammar, contra GROUND.

Now, I do not know why anyone would accept (16). And I do not know how one might argue for (16). For suppose N is the purportedly most natural neural relation such that $\mathsf{T}[N, s, G]$ and $N(s, G)$, where G is the correct grammar for s's language. Why think that N must be more natural than each permuted cousin N_i^*? One might argue that each N_i^* is partially defined in terms of N, such that the metaphysical definition of N_i^* in fundamental terms is longer or more complex than that of N, which entails that N is the more natural relation. But of course N is also definable in terms of N_i^*, as follows: $N(x, G)$ iff $N_i^*(x, i^{-1}(G))$. And even though *we* defined the relational predicate 'N_i^*' in terms of N, it is not clear why we should expect N_i^*'s metaphysical definition to look anything like *our* definition of 'N_i^*'. So it is not clear that we have not already encountered counterexamples to (16).

Moreover, there is a speculative metaphysical argument that (16) is false. There seems to be a way to construct, for any purportedly most natural neural relation realizing T in s and relating s to the correct grammar for their language, a *more natural* relation relating s to an incorrect grammar for their language that also realizes T in s. To see this, suppose $\mathsf{T}[\mathbf{N}, \mathbf{s}, \mathbf{G}]$ and $\mathbf{N}(\mathbf{s}, \mathbf{G})$. And suppose $\mathbf{N}$ is a neural relation that is the most natural relation realizing T in $\mathbf{s}$. Given the truth of physicalism and the Chomskyan hypothesis that $\mathbf{N}$ is a computational relation, $\mathbf{N}(\mathbf{s}, \mathbf{G})$ supervenes on $\mathbf{s}$'s having some more specific, more fundamental, non-computational, neurobiological property $\mathbf{B}$.[51] $\mathbf{B}$ is less disjunctive, less general, less determinable, less gerrymandered, and so *more natural* than $\mathbf{N}$. And $\mathbf{B}(\mathbf{s})$ in turn supervenes on $\mathbf{s}$'s having a more natural biochemical property $\mathbf{C}$, which supervenes on $\mathbf{s}$'s having a yet more natural physical property $\mathbf{P}$.

Simplifying massively, let *nat* be a function from a property to a naturalness score between 0 (maximally grueish) and 1 (perfectly natural). And suppose the scores of $\mathbf{N}$ and its subvenient properties are:

$$nat(\mathbf{N}) = 0.3,\ nat(\mathbf{B}) = 0.4,\ nat(\mathbf{C}) = 0.5,\ nat(\mathbf{P}) = 0.6$$

Next, define $\mathbf{N}^*$ as before: $\mathbf{N}^*(x, G)$ iff $\mathbf{N}(x, i(G))$. And let us suppose that the $nat(\mathbf{N}^*) = 0.2$. $\mathbf{N}^*$ would be tied with $\mathbf{N}$, as it were, except 0.1 is deducted because it is defined in terms of both $\mathbf{N}$ and our bijection i.

Now, even though $\mathbf{N}$ is more natural than $\mathbf{N}^*$, it seems that $\mathbf{N}$ is less natural than the following relation:

$$\mathbf{N}^{\mathbf{C}}(x, G) \text{ iff } G = i(\mathbf{G}) \text{ and } \mathbf{C}(x)$$

where i is defined as before, but let us assume that $i(\mathbf{G})$ is an incorrect grammar for $\mathbf{s}$'s language. Given that $nat(\mathbf{C}) = 0.5$, the score of $\mathbf{N}^{\mathbf{C}}$ should be something like $0.5 - 0.1$ for being defined in terms of i. But if so, then $nat(\mathbf{N}^{\mathbf{C}}) = 0.4$. And so $\mathbf{N}^{\mathbf{C}}$ is more natural than $\mathbf{N}$.[52]

Moreover, given that $\mathbf{N}$ realizes T in $\mathbf{s}$, I think $\mathbf{N}^{\mathbf{C}}$ does as well. This is because $\mathbf{C}(\mathbf{s})$ and $\mathbf{N}^{\mathbf{C}}(\mathbf{s}, i(\mathbf{G}))$ are logically equivalent. And so because $\mathbf{C}(\mathbf{s})$ entails $\mathbf{N}(\mathbf{s}, \mathbf{G})$, so does $\mathbf{N}^{\mathbf{C}}(\mathbf{s}, i(\mathbf{G}))$. $\mathbf{N}(\mathbf{s}, \mathbf{G})$ supervenes upon $\mathbf{N}^{\mathbf{C}}(\mathbf{s}, i(\mathbf{G}))$. And so, arguably, whatever causal relations the $\mathbf{N}(\mathbf{s}, \mathbf{G})$-state stands in, the $\mathbf{N}^{\mathbf{C}}(\mathbf{s}, i(\mathbf{G}))$-state also stands in. And so, given $\mathsf{T}[\mathbf{N}, \mathbf{s}, \mathbf{G}]$, I think we also have $\mathsf{T}[\mathbf{N}^{\mathbf{C}}, \mathbf{s}, i(\mathbf{G})]$. To see this, consider again what $\mathsf{T}[\mathbf{N}, \mathbf{s}, \mathbf{G}]$ says:

(17) If it were the case that $\mathbf{N}(\mathbf{s}, \mathbf{G})$ and ..., then, *ceteris paribus*:

(i) for any S and m, if $\mathbf{G}$ generates S and assigns m to S, then: if $\mathbf{s}$ were to perceive S, then $\mathbf{s}$ would interpret S as meaning m ……

Given that $\mathbf{N}(\mathbf{s}, \mathbf{G})$ supervenes on $\mathbf{C}(\mathbf{s})$, it follows that:

(18) If it were the case that $\mathbf{N}^{\mathbf{C}}(\mathbf{s}, i(\mathbf{G}))$ and …, then, *ceteris paribus*:
(i) for any S and m, if $\mathbf{G}$ generates S and assigns m to S, then: if $\mathbf{s}$ were to perceive S, then $\mathbf{s}$ would interpret S as meaning m ……

For whatever is causally or counterfactually dependent on $\mathbf{N}(\mathbf{s}, \mathbf{G})$ will be likewise dependent on $\mathbf{N}^{\mathbf{C}}(\mathbf{s}, i(\mathbf{G}))$. This is because the causal powers of the higher-level neural $\mathbf{N}(\mathbf{s}, \mathbf{G})$-state will be identical to causal powers of the lower-level chemical $\mathbf{N}^{\mathbf{C}}(\mathbf{s}, i(\mathbf{G}))$-state. For, given physicalism, the neural state will have no *sui generis*, emergent causal power that is not shared with or inherited from its underlying chemical state.[53]

Now, with (18) in hand, because $\mathbf{G}$ and $i(\mathbf{G})$ are practically equivalent, we can, as before, replace '$\mathbf{G}$' with '$i(\mathbf{G})$' in (18) to arrive at $\mathsf{T}[\mathbf{N}^{\mathbf{C}}, \mathbf{s}, i(\mathbf{G})]$. Thus, $\mathbf{N}$ cannot be the most natural relation realizing T in $\mathbf{s}$. $\mathbf{N}^{\mathbf{C}}$ is more natural and also realizes T in $\mathbf{s}$. But, unlike $\mathbf{N}$, $\mathbf{N}^{\mathbf{C}}$ relates $\mathbf{s}$ to an incorrect grammar for their language. And this means that (16) is false.

For these reasons, I am dubious of (16). And (16) is unmotivated anyways. And so I do not think that FUNCTIONALISM^{+} is a promising solution.

7 | CONCLUDING THOUGHTS

At present, I do not see how to reconcile FUNCTIONALISM and GROUND. But, conditional on the reality of cognizing, FUNCTIONALISM looks like the right metaphysics of it. The Chomskyan can of course deny FUNCTIONALISM. But, as I explained above, it is not as if there is any other worked out account to replace it. They could also say that cognizing is a primitive or emergent psychological relation. Or that it is not realized by, but is rather *type-identical to*, some neural relation. Or they could adopt 'mysterianism'.[54] All of these positions deserve further discussion elsewhere. But none strike me as attractive.

So I am inclined to conclude that GROUND is false. If so, what this means is that, despite many philosophers' suggestions to the contrary, we cannot safely appeal to cognizing to deal with semantic or metasemantic indeterminacy or underdeterminacy. Cognizing cannot plug the hole left in one's favorite theory of meaning after it is pointed out that it posits a metasemantic base that fails to fix the meanings of large swathes of language. Appeals to cognizing as a plug are made, for example, in reaction to well-established shortcomings of psychological or 'intention-based' theories of meaning in the traditions of Grice (1957) and Lewis (1969, 1975). These theories analyze a sentence's meaning in terms of the contents of its users' communicative intentions. And due to their 'sentence-first' orientation, they face systematic problems in accounting for subsentential meaning and semantic determinacy beyond language's used or usable fragment.[55] Given that we only use a tiny fragment of our language, how can use determine that the grammar of our language is our preferred 'straight' one rather than a practically equivalent 'bent' one?

But the same problem arises for any theory of meaning. The problem is, in a word, compositionality. How do we explain it? The best strategy *seems* to be to supplement one's metasemantics with cognizing, thereby verifying GROUND.[56] For if we do cognize grammars, this move just widens the psychological metasemantic base.

In my view, this strategy looks unworkable. For if we cognize a straight grammar, we must also cognize a plurality of bent grammars. We cannot rely on cognizing to fix even the syntax and phonology of our language, let alone its semantics. Language just ain't in the head.

ENDNOTES

[1] At least as I read him. Others might suspect that I have already fatally confused competence with performance in taking competence to be that which *enables* performance. But this has long been his expressed view: see Chomsky 1964 (pp. 7–27), 1965 (pp. 3–9), 1966 (p. 75, n. 2), 1968 (pp. 27–31), 1972 (p. 11, 26, 103–4), Chomsky and Katz 1974 (pp. 365–67), Chomsky 1975a (pp. 304–6), 1975b (pp. 12–13, 164, 195–98), Chomsky and Lasnik 1977 (pp. 427–28), Chomsky 1980 (pp. 59–62), 1982a (pp. 429, 431), 1986 (pp. 22, n. 10; 24), 1992 (pp. 211–12), 1994 (p. 159), 1995a (p. 12), 1995b (pp. 12–16), 2000 (pp. 9–10), 2005 (pp. 3–4), and 2012 (p. 69).

[2] "By 'language' I mean 'human language'" (Chomsky 1994, p. 155).

[3] The LF of *S* is a structure representing exactly the syntactic information about *S* relevant to its semantic interpretation; it would be better to call it 'semantic form' (Szábo 2012, p. 105). For helpful philosophical discussions of LF, see Neale 1993, 1994, Ludlow 2002, and King 2002.

[4] In Chomsky 1995b, 'PF' means phonetic form, but elsewhere often means phonological form. Scheer 2010 argues that Chomsky's usage is equivocal (pp. 616–18). Standardly, *e*'s phonological form represents the structure of the sound associated with *e* as mentally represented by the speaker, whereas *e*'s phonetic form represents its pronunciation (Myers 2000, pp. 245–46). But, as Carr 2012 argues, there "is no consensus in the phonological literature as to whether it is possible to adopt a clear distinction (or indeed, any distinction) between phonetics and phonology" (p. 403). See also Bromberger and Halle 1986 (pp. 139–43), 1989 (pp. 51–3), 2000 (pp. 17–21, 30–7); Bromberger 2012 (pp. 83, 88–92); and Carr 2012 (pp. 403–12).

[5] As far as I know, I am the first to explicitly advocate this. Perhaps this is because philosophers like to think of an expression as a "finite sequence of types of vocal sounds or types of marks," following Lewis 1969 (p. 142) and Quine 1960 (pp. 194–95). But neither argue for this; they stipulate it. And by now this model of expressions is quite unscientific. So let us follow linguists and instead model expressions as tuples of structures. I admit how to individuate words may be a book-keeping issue. But by keeping them this way, we put the actual-language relation to work in metasyntax, metaphonology, and metasemantics, and unify these enterprises.

[6] But Ludlow takes the physical subvenient state to be widely not narrowly individuated (pp. 117–18, 140).

[7] See also Chomsky 2002 for an extended discussion.

[8] Perhaps an exception is the account developed by Evans 1981, Davies 1987, and Peacocke 1989. But even they are less than explicit about what their account entails regarding the metaphysical status of cognizing vis-à-vis the brain. I address this view in section 6.3 below.

[9] Collins 2004 rejects this, and any causal-functional theory of cognizing, by arguing that cognizing does "not enter into the causal nexus leading to linguistic acts" (p. 527). (George 1989 (pp. 92, 98–9) and Chomsky 1986 (pp. 238–43) raise similar worries.) But cognizing can be definitively causally characterized even if it does not *cause* anything. After all, cognizing is a state, and perhaps only events cause. But a state of cognizing nevertheless *causally enables* mental processes like parsings to unfold, and enables causal transitions between mental events, such as transitions from events of linguistic perception to events of understanding (see fn. 1).

[10] To simplify, I focus only on the 'forward-looking' part of the cognizing-role that specifies that which causally depends on cognizing. But everything I say can be reworked to include the 'backward-looking' part that specifies whatever cognizing as such causally depends on. This part may, for example, specify the causes of cognizing and language acquisition.

[11] In the terms of Hauser et al. 2002, the suggestion is that '...' will say that *s*'s cognizing *G*—*s*'s "FLN" or "faculty of language" in the "narrow sense"—"interfaces" with *s*'s "FLB" or "faculty of language" in the "broad sense"—*s*'s "sensory-motor" and "conceptual-intensional" systems, the neural mechanisms responsible for communication and thought, respectively (pp. 1570–71).

[12] Pietroski and Rey 1995 cite Chomskyan theories of competence as issuing in substantial, *ceteris paribus* laws in defense of the explanatory power and non-vacuity of such laws (pp. 83, 107).

[13] See also Pollock 1989 for an extended argument that a true functional theory of any "organism" cannot be interpreted as making exceptionless generalizations about how it actually functions (pp. 52–61, 115–16).

[14] Here I depart from Lewis's 1970, 1972 formulation of functionalism, following Harman 1973 (pp. 41–3), Field 1978 (pp. 46–8, fn. 19), and Shoemaker 1981 (pp. 96–97) in taking T as a universal generalization, and in defining a realization of T as a relation figuring in a relation-speaker-grammar triple satisfying $\mathsf{T}[R, s, G]$, rather than a relation satisfying $\forall s \forall G.\mathsf{T}[R, s, G]$. For more on related subtleties, see Loar 1981 (pp. 48–56) and David 1997 (pp. 144–48).

[15] See Lewis 1972 (pp. 256–7) for this argument as applied to the notions of folk psychology, and see Horgan 1984 (pp. 31–5) and Byrne 1999 (pp. 353–56) for discussion. See Rey 1997 (pp. 172–88, 203–4) for the argument as applied to the notions of scientific psychology.

[16] Another argument for FUNCTIONALISM is metaphysical. Plausibly, cognizing supervenes on the physical; there can be no difference in who cognizes which grammar without a physical difference. This requires an explanation (Horgan 1993, Wilson 2005). And the best explanation is that cognizing is functionally realized (Melnyk 2003 (p. 26, 55–7), 2018; Tiehen 2016). For as Tiehen 2018 argues, functional realization's only serious contender is 'grounding', which is arguably not a strong contender (Wilson 2014, 2018 and Melnyk 2016).

[17] Compare Field's 1978 argument that functionalism about the belief relation leaves unexplained how it is possible for believers to stand to a proposition in whatever physical relation plays the belief-role (pp. 30–33). See also Field 2001 (pp. 69–71) and Stalnaker 1984 (pp. 14–18) for discussion.

[18] This is evident in the Pietroski, Hornstein, and Chomsky passages cited above. And it is claimed or heavily implied in Marr 1982 (pp. 28–29, 357); Stabler 1983; Boden 1984 (pp. 26–8); Egan 2003; Rey 2003 (pp. 120–23); Collins 2000 (pp. 469–70), 2004 (pp. 525, 529–30), 2007 (pp. 634–36); Devitt 2006 (pp. 66–71); Hornstein and Pietroski 2009 (pp. 114, 123); Pietroski 2010 (pp. 250–52), 2011 (pp. 473–75; 474, fn. 4); Johnson 2014 (pp. 52–3); Schiffer 2015 (pp. 62–4); Berwick and Chomsky 2016 (pp. 128–39); Poeppel 2017 (pp. 156–63); and Chomsky 2018 (pp. 34–35). This is but a mere sampling, for this view is advanced implicitly anywhere cognizing is said to be posited by a true computational psychology.

[19] We might want to add a third component,

> (iii) necessarily, if some system in s's brain implements G, then there is some neural property F such that s has F, and, s's having F entails that some mechanism in s's brain implements G.

to ensure that $N(s, G)$ supervenes narrowly on s's brain state. But I will omit this and remain neutral about whether implementing a grammar is a 'wide' or 'narrow' state. But note that if it *is* wide, then N is not a purely neural relation.

[20] This will not be easy, as I argue in section 6.2 below.

[21] Adherents of psychofunctionalism include Putnam 1960; Fodor 1968; Harman 1973; Field 1978; Loar 1981 (pp. 75–80); Lycan 1981, 1987, 1996; Kim 1999 (pp. 10–11), 2005 (although not for "qualia", pp. 26–27); Melnyk 2003; Rey 1997 (pp. 187–204), 2020 (p. 343, fn. 10); Piccinini 2020 (pp. 96–106, 322); and many others. As Polger 2004 remarks, "most functionalists [...] are psychofunctionalists" (p. 96, fn. 28).

[22] For example, of the five objections to psychofunctionalism discussed by Hill 1991 (pp. 69–82), the first, fourth, and fifth are of type-(b), the second and third are of type-(a) and type-(b), and all five cannot be (straightforwardly) transposed into objections to FUNCTIONALISM for these reasons. Also, Shoemaker's 1981 (pp. 112–14) argument that psychofunctionalism is "obviously wrong" is also an objection of type-(b). And so is Block's 1978 objection that psychofunctionalism is "hopelessly chauvinistic" (pp. 310–14) (see Clark 1986 for discussion of Shoemaker and Block).

[23] See Papineau 1996, Schurz 2005, and Andreas 2021 for more on the Ramsey-Carnap-Lewis account of theoretical terms.

[24] I draw inspiration from an argument in Field 1986 (p. 108) against Stalnaker's 1984 version of functionalism about belief and desire, from a similar argument against generalized forms of functionalism in Bealer 1984, 1997 (pp. 106–17), from some anti-Chomskyan moves made in Quine 1972 and Lewis 1975, and from some cryptic remarks in Kripke 1982 (p. 30, fn. 22; p. 97, fn. 77).

[25] For Lewis 1975, sententially equivalent grammars are grammars that "[generate] the same language" (p. 20), for his 'languages', unlike mine, only map sentences to meanings. A similar notion of sentential equivalence—but of semantic theories—appears in Williams 2008.

[26] Here and in what follows I display T's causal claims as counterfactuals. But nothing will ultimately turn on this. I just find (2) easier to parse than 'For any speaker s and grammar G, if s were to cognize G and ..., then, *ceteris paribus*, (i) the state of s's cognizing G would causally enable any event of s's perceiving a sentence S generated by G to cause an event of s's interpreting S as having the meaning that G assigns to S'.

[27] I assume that if a counterfactual '$(A \wedge B) > C$' is true and 'A' is definitionally or analytically equivalent to 'D', then '$(D \wedge B) > C$' must also be true. This strikes me as nonnegotiable, even though a hyperintensional theory of counterfactuals might falsify the version of this principle invoking *intensional* equivalence. And it also seems clear enough that definitionally equivalent token states must stand in exactly the same causal relations (cf. Bealer 1984, pp. 295–96). So the move from (3) to (4) would go through even if they were converted from counterfactuals into explicit statements of causal relatedness.

[28] One might object that COMPUTATIONAL HYPOTHESIS has not been shown to be false because there is no guarantee that whenever N is a neural computational relation, each N_i^* is also a *neural* computational relation; although the latter may be defined in terms of N, the 'neuralness' of N might not transfer over such a definition. But I am unsure whether such a fine-grained criterion for a relation's being neural can be sustained in a principled way. And even if it can, although this move might save COMPUTATIONAL HYPOTHESIS, it can only save FUNCTIONALISM when reformulated as the stronger thesis that, necessarily, x cognizes G just if there is a unique *neural* relation R such that $\mathsf{T}[R, x, G]$ and $R(x, G)$. But why think it is metaphysically impossible for cognizing to be non-neurally realized, even if, as a matter of fact or law (or even metaphysical necessity) cognizing cannot be non-neurally realized in humans? Moreover, might not non-neural implants one day assist human brains to realize cognizing? (Thanks to Stephen Schiffer for discussion here.)

[29] See Lewis 1994 (p. 299–301) and also Braddon-Mitchell and Jackson 2007 (pp. 52–59).

[30] See Stich 1971, 1978 for some widely-discussed arguments in this direction, against the claim that we have ordinary knowledge of grammatical axioms and rules on the basis of which we might derive knowledge of their theorems. Graves et al. 1973 takes issue with Stich's arguments, but the view they defend is that a speaker has "knowledge that a particular expression has a certain grammatical property [...] only upon encountering that expression" (p. 320); thus, they do not defend (9), but rather some presumably *ceteris paribus* generalization about performance (see my discussion of (9*) below). For recent discussion of these issues, see Collins 2022.

[31] See Chomsky 1975b (pp. 164–66), 1980 (pp. 69–70), and 1986 (pp. 265–66).

[32] If this is true, this opens the door to an alternative solution, inspired by the view of cognizing defended in Knowles 2000. For we might say that T includes (13) instead of (9):

(13) If s were to cognize G, then, for any grammatical principle (rule or axiom) P of G, s would know P.

Knowles argues, in effect, that (13) is true. But I think this is a non-solution for the same reasons that including (9) in T is. For critical discussion of Knowles's view, see Rattan 2002 and Collins 2004 (pp. 17–19, fn. 12). See also Waights Hickman 2021 for a recent defense of a view similar to Knowles's.

[33] See Uriagereka 2012 for informed speculation in this direction (pp. 2–12). And see Langendoen and Postal 1984 for related arguments that natural languages are so large that they cannot be generated by computable or "constructive" grammars (although they suggest that a speaker might implement a grammar for a "computable subset" of their language (pp. 97–103)).

[34] Causal structuralism is defended in Chalmers 1994, 1996, 2012, and variants of it in Copeland 1996, Dresner 2010, Egan 1992, Godfrey-Smith 2009, and Scheutz 2001. For a discussion of this view in application to grammar-implementation, see Ludlow 2011 (pp. 119–20).

[35] In more detail, let us suppose that G is my grammar, and a minimalist grammar of the kind formalized in Collins and Stabler 2016. The lexicon of G will be a finite set of lexical items: triples of sets of semantic, syntactic, and phonological 'features', ⟨SEM, SYN, PHON⟩, drawn from universal sets of such features SEM-F, SYN-F, and PHON-F that are part of Universal Grammar, as aspect of our innate psychobiological endowment. Surely, many of these features will be arbitrary in the Saussurean sense, such that our languages could have worked the same if, say, some elements of PHON-F had been different. (For a sophisticated discussion of this issue, see Bierwisch 2014.) Thus, it is plausible that my language could have had the same derivational structure even if some of the phonological features that make up its lexical items (drawn from PHON-F) were different. But that is just to say that there are many grammars distinct from, but with the same derivational structure as, G.

[36] See Haselanger et al. 2003 (pp. 8–12, 16–19) for an interesting discussion of issues in this connection.

[37] Take Rescorla's 2014 account, *computational descriptivism*, on which causal-structural isomorphism is necessary but not sufficient for implementation. On this view, "physical system P realizes/implements computational

model *M* just in case computational model *M* accurately describes" *P*, or just if "*P* reliably moves through 'state space' according to mechanical instructions encoded by *M*" (p. 1278–80). If grammars can be treated as models, say that *P* implements *G* just if *P* is accurately described by *G*. This view also seems to overgenerate. To be accurately described by a model is to reliably "[conform] to [the] instructions" encoded in that model", which "requires *doing what the instructions say*" (p. 1288). So, if *G* accurately describes *P*, *P* must have "an ability to instantiate states specified by" *G* in accordance with its instructions. This, for Rescorla, is the key difference between a mechanism's causal organization being isomorphic to a model and its implementing it: having the capacity to do what it instructs. Now, the 'instructions' of a grammar are its *rules*, not its axioms or lexical entries. So if *P* is accurately described by *G*, then *P* will also be accurately described by a lexically permuted G^*. So if *P* implements one, *P* implements both, which leads to the same bad result. But perhaps Rescorla's view can avoid this result if grammars are modeled in the right way. Perhaps *G* qua model will induce a state space description that includes states corresponding to *G*'s lexical entries. If so, then for *P* to implement the *G*, *P* must be able to enter those states. If so, there is no guarantee that *P* will also be able to enter the states of G^*, and no guarantee that *P* is accurately described by G^* if by *G*. But *can* grammars be accurately modeled like this? The answer, Rescorla suggests, is not entirely objective and depends "on our pragmatic or explanatory goals" (p. 1279). But if so, there is no non-interest-relative fact of the matter about whether his view can avoid the problem.

[38] For a defense of the possibility of a speaker like Floyd, see Schiffer 1993 (pp. 236–39), 2006 (pp. 280-82), in response to Lewis's 1992b claim that, even if a bent grammar were "written into the brain", a speaker's language would nevertheless be one determined by a straight grammar (p. 151, fn. 6).

[39] I assume, as seems plausible, that in modifying the bent rules of G^* to get G^{**}, swapping out the assignment of the semantic value of 'the cat' to descriptions 100,000-words or longer for an assignment to them of the semantic value of 'the mouse' does not induce a change in derivational structure.

[40] See also Davies 1989, 1995 and Peacocke 1986. This account has been endorsed more recently in Pereplyotchik 2017 (pp. 169–80) and Rey 2020 (pp. 113–16).

[41] In what follows I only assess Peacocke's formulation of this view, trusting his assessment that his account "can be shown to be equivalent" (p. 182, fn. 2) to Davies's 1987, and Davies's 1989 assessment that "the two accounts turn out to be equivalent" (p. 132).

[42] As above, I transpose talk of causal powers into counterfactual talk, for simplicity.

[43] For further problems with the Evans-Davies-Peacocke view, see Barber 2007.

[44] The contemporary notion of naturalness is originally due to Merrill 1980. See Dorr and Hawthorne 2013 for critical discussion.

[45] For simplicity, let us ignore the possibility of naturalness ties.

[46] See Lewis 1969 (pp. 198–200), Lewis 1975 (pp. 18–21), and Lewis 1992b (pp. 109–110; fn. 6). And for discussion see Cresswell 1994 (pp. 98–114), Weatherson 2013, Schwarz 2014 (pp. 18–21), Williams 2015 (pp. 370–72), and Janssen-Lauret and MacBride 2020 (pp. 4–6).

[47] Lewis also considers a Chomskyan view of grammar-evaluation, on which the best grammar for a speaker is the one they cognize (p. 199). But there is no such grammar as *the* grammar a speaker cognizes.

[48] This is proven in Michaelis 2001. Two grammars are strongly equivalent just if they generate the same strings and assign them the same structural descriptions or tree-like syntactic structures; see Chomsky 1963.

[49] In reply, the Chomskyan might object that if the neural evidence can confirm that a speaker's language is of G^* and not *G*, this must be because some neural fact ϕ about their brain *entails* that G^* is *the* grammar of their language. And if so, or so they might say, then ϕ-type facts must neurally constitute cognizing without issue. But this is a mistake. Compare the case of consciousness. The neural evidence confirms that other humans are conscious, given what we know about consciousness's neural correlates, which confirms that this is possible. And we can accept this without thinking that the neural facts *entail* that others are conscious. Someone who accepts the possibility of zombies, for instance, could still think the neural evidence supports the existence of other consciousnesses. Likewise, there might be neural evidence about the character of a speaker's language or grammar even if their brain does not fix their language or grammar.

[50] A thesis like this is supported by a more complicated metasemantic argument based on a naturalness-based version of the Ramsey-Carnap-Lewis account of the implicit definition of theoretical terms, like the one recommended by Hawthorne 1994.

[51] Here I assume that the property $\lambda x.\mathbf{N}(x, \mathbf{G})$ supervenes on narrow physical properties of the brain/body. But this assumption is inessential to the argument.

[52] Perhaps another 0.1 should be deducted because the definition of $\mathbf{N}^{\mathbf{C}}$ mentions the grammar **G**. If so, we can instead use the relation defined as follows: $\mathbf{N}^{\mathbf{P}}(x, G)$ iff $G = i(\mathbf{G})$ and $\mathbf{P}(x)$. And its score will then be 0.4.

[53] In support of this, see Wilson's 2021 argument that this constraint on causal powers is satisfied by all candidate physicalistic accounts of how higher-level goings-on supervene or depend upon lower-level goings-on (pp. 55–69).

[54] Perhaps Chomsky 2009 can be read as leaving open such a position.

[55] This is the 'meaning without use' problem. See Hawthorne 1990, 1993; Lewis 1992b, 1992a; Schiffer 1993 (pp. 235–39), 2006 (pp. 278–82), 2017 (pp. 54–55, 63–57); Ray 1995; and Keiser 2021.

[56] This is proposed by Loar 1976 (pp. 157–61), 1981 (pp. 257–60); Davies 2000 (pp. 52–56); and Schiffer 2006 (pp. 285–87). (Plus, Lewis 1992a reveals this to be his fall-back position.) And Laurence 1996, Yalcin 2014, and Pietroski 2018 (pp. 58–60) read as favorable to the proposal. See also Keiser 2021 (pp. 20–26) for more recent discussion.

REFERENCES

Andreas, Holger (2021). "Theoretical terms in science". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Stanford University, Fall 2021 Edition, URL https://plato.stanford.edu/archives/fall2021/entries/theoretical-terms-science/.

Balcarras, David (2023). "Is meaning cognized?" *Mind and Language*, pp. 1–20. 10.1111/mila.12452.

Barber, Alex (2007). "Linguistic Structure and the Brain". *Croatian Journal of Philosophy*, 21, pp. 317–341.

Bealer, George (1984). "Mind and Anti-mind: Why Thinking Has No Functional Definition". *Midwest Studies in Philosophy*, 9(1), pp. 283–328.

——— (1997). "Self-Consciousness". *The Philosophical Review*, 106(1), pp. 69–117.

Berwick, Robert C. and Noam Chomsky (2016). *Why Only Us: Language and Evolution*. The MIT Press.

Bierwisch, Manfred (2014). "The role of arbitrariness from a minimalist point of view". In: *Minimalism and Beyond: Radicalizing the Interfaces*. Ed. by Peter Kosta, Steven L. Franks, Teodora Radeva-Bork, and Lilia Schürcks. John Benjamins, pp. 392–416.

Block, Ned (1978). "Troubles with Functionalism". In: *Perception and Cognition: Issues in the Philosophy of Science*. Ed. by C. W. Savage. University of Minnesota Press, pp. 261–325.

Boden, Margaret A. (1984). "What is Computational Psychology? (I)". *Proceedings of the Aristotelian Society*, 58, pp. 17–35.

Braddon-Mitchell, David and Frank Jackson (2007). *The Philosophy of Mind and Cognition*. Blackwell, Second Edition.

Bromberger, Sylvain (2012). "Vagueness, Ambiguity, and the "Sound" of Meaning". In: *Analysis and Interpretation in the Exact Sciences: Essays in Honour of William Demopoulos*. Ed. by Melanie Frappier, Derek Brown, and Robert DiSalle. Springer, pp. 75–93.

Bromberger, Sylvain and Morris Halle. (1986). "On the Relationship of Phonology and Phonetics: Comments on B. Lindblom 'on the Origin and Purpose of Discreteness and Invariance in Sound Patterns'". In: *Invariance and Variability in Speech Processes*. Ed. by J. S. Perkell and D. H. Klatt. L. Erlbaum Associates, pp. 510–520.

——— (1989). "Why phonology is different". *Linguistic Inquiry*, 20(1), pp. 51–70.

——— (2000). "The Ontology of Phonology (Revised)". In: *Phonological Knowledge: Conceptual and Empirical Issues*. Ed. by Noel Burton-Roberts, Philip Carr, and Gerard Docherty. Oxford University Press, pp. 19–37.

Byrne, Alex (1999). "Cosmic Hermeneutics". *Philosophical Perspectives*, 13, pp. 347–383.

Carnap, Rudolf (1961). "On the Use of Hilbert's Epsilon-Operator in Scientific Theories". In: *Essays on the Foundations of Mathematics*. Ed. by Yehoshua Bar-Hillel. Magnes Press, pp. 156–164.

Carr, Philip (2012). "The Philosophy of Phonology". In *Philosophy of Linguistics*. Ed. by Ruth Kempson, Tim Fernando, and Nicholas Asher. Elsevier, pp. 403–444.

Chalmers, David J. (1994). "On implementing a computation". *Minds and Machines*, 4(4), pp. 391–402.

——— (1996). "Does a Rock Implement Every Finite-State Automaton?" *Synthese*, 108(3), pp. 309–333.

——— (2012). "The Varieties of Computation: A Reply". *Journal of Cognitive Science*, 3, pp. 211–248.

Chomsky, Noam (1963). "Formal Properties of Grammars". In: *Handbook of Mathematical Psychology, Volume II*. Ed. by R. Duncan Luce, Robert R. Bush, and Eugene Galanter. Wiley, pp. 323–418.

——— (1964). *Current Issues in Linguistic Theory*. Mouton.

——— (1965). *Aspects of the Theory of Syntax*. The MIT Press.

——— (1966). *Cartesian Linguistics: A Chapter in the History of Rationalist Thought*. Cambridge University Press, Third Edition, 2009.
——— (1968). *Language and Mind*. Harcourt Brace Jovanovich, Inc.
——— (1972). *Studies on Semantics in Generative Grammar*. Mouton.
——— (1975a). "Knowledge of language". In: *Language, Mind, and Knowledge*. Ed. by Keith Gunderson. Minnesota Studies in the Philosophy of Science, Vol. VII, University of Minnesota Press, pp. 299–320.
——— (1975b). *Reflections on Language*. Pantheon.
——— (1980). *Rules and Representations*. Columbia University Press.
——— (1982a). "A Note on the Creative Aspect of Language Use". *The Philosophical Review*, 91(3), pp. 423–434.
——— (1982b). "Mental Representations". *Syracuse Scholar*, 4(2), pp. 5–21.
——— (1986). *Knowledge of Language: Its Nature, Origin, and Use*. Praeger.
——— (1988). *Language and Problems of Knowledge: The Managua Lectures*. The MIT Press.
——— (1992). "Explaining Language Use". *Philosophical Topics*, 20(1), pp. 205–231.
——— (1994). "Noam Chomsky". In: *A Companion to the Philosophy of Mind*. Ed. by Samuel D. Guttenplan. Blackwell, pp. 153–167.
——— (1995a). "Language and Nature". *Mind*, 104(413), pp. 1–61.
——— (1995b). *The Minimalist Program*. The MIT Press.
——— (2000). *New Horizons in the Study of Language and Mind*. Cambridge University Press.
——— (2002). "Language and the brain". In: *On Nature and Language*. Ed. by Adriana Belletti and Luigi Rizzi. Cambridge University Press, pp. 61–91.
——— (2005). "Three Factors in Language Design". *Linguistic Inquiry*, 36(1), pp. 1–22.
——— (2009). "The Mysteries of Nature: How Deeply Hidden?" *The Journal of Philosophy*, 106(4), pp. 16–200.
——— (2012). *The Science of Language: Interviews with James McGilvray*. Cambridge University Press.
——— (2018). "Mentality Beyond Consciousness". In: *Ted Honderich on Consciousness, Determinism, and Humanity*. Ed. by Gregg D. Caruso. Springer, pp. 33–46.
Chomsky, Noam and Morris Halle (1968). *The Sound Pattern of English*. Harper & Row.
Chomsky, Noam and Jerrold J. Katz (1974). "What the Linguist is Talking About". *The Journal of Philosophy*, 71(12), pp. 347–367.
Chomsky, Noam and Howard Lasnik (1977). "Filters and Control". *Linguistic Inquiry*, 8(3), pp. 425–504.
Clark, Austen (1986). "Psychofunctionalism and Chauvinism". *Philosophy of Science*, 53(4), pp. 535–559.
Collins, Chris and Edward Stabler (2016). "A Formalization of Minimalist Syntax". *Syntax*, 19(1), pp. 43–78.
Collins, John (2000). "Theory of Mind, Logical form and Eliminativism". *Philosophical Psychology*, 13(4), pp. 465–490.
——— (2004). "Faculty Disputes". *Mind and Language*, 19(5), pp. 503–533.
——— (2007). "Review of *Ignorance of Language* by Michael Devitt". *Mind*, 116(462), pp. 416–423.
——— (2022). "Chomsky, Cognizing, and Tacit Knowledge". In: *The Routledge Handbook of Philosophy and Implicit Cognition*. Ed. by J. Robert Thompson. Routledge, pp. 237–246.
Copeland, B. Jack (1996). "What is computation?" *Synthese*, 108, pp. 335–359.
Cresswell, Maxwell J. (1994). *Language in the world: A philosophical enquiry*. Cambridge University Press.
David, Marian (1997). "Kim's Functionalism". *Philosophical Perspectives*, 11, pp. 133–148.
Davies, Martin (1987). "Tacit Knowledge and Semantic Theory: Can a Five per cent Difference Matter?" *Mind*, 96(384), pp. 441–462.
——— (1989)."Tacit Knowledge and Subdoxastic States." In: *Reflections on Chomsky*. Ed. by Alexander George. Blackwell, pp. 131–152.
——— (1995). "Two Notions of Implicit Rules". *Philosophical Perspectives*, 9, pp. 153–183.
——— (2000). "Persons and their underpinnings". *Philosophical Explorations*, 3(1), pp. 43–62.
Devitt, Michael (2006). *Ignorance of Language*. Oxford University Press.
Dorr, Cian and John Hawthorne (2013). "Naturalness". In: *Oxford Studies in Metaphysics, Volume 8*. Ed. by Karen Bennett and Dean W. Zimmerman. Oxford University Press, pp. 3–77.
Dresner, Eli (2010). "Measurement-Theoretic Representation and Computation-Theoretic Realization". *The Journal of Philosophy*, 107(6), pp. 275–292.
Egan, Frances (1992). "Individualism, Computation, and Perceptual Content". *Mind*, 101(403), pp. 443–459.

——— (2003). “Naturalistic Inquiry: Where does Mental Representation Fit in?” In: *Chomsky and His Critics*. Ed. by Louise M. Antony and Norbert Hornstein. Blackwell, pp. 89–104.

Evans, Gareth (1981). “Semantic Theory and Tacit Knowledge”. In: *Wittgenstein: To Follow a Rule*. Ed. by Steven H. Holtzman and Christopher M. Leich. Routledge, pp. 118–137.

Field, Hartry (1978). “Mental Representation”. *Erkenntnis*, 13(1). pp. 9–61.

——— (1986). “Stalnaker on intentionality”. *Pacific Philosophical Quarterly*, 67, pp. 98–112.

——— (2001). “Postscript [to “Mental Representation”]”. In *Truth and the Absence of Fact*, pp. 68–82. Oxford University Press.

Fodor, Jerry A. (1968). *Psychological Explanation: An Introduction to the Philosophy of Psychology*. Random House.

George, Alexander (1989). “How not to become Confused about Linguistics”. In: *Reflections on Chomsky*. Ed. by Alexander George. Blackwell, pp. 90–110.

Godfrey-Smith, Peter (2009). “Triviality arguments against functionalism”. *Philosophical Studies*, 145, pp. 273–295.

Graves, Christina, Jerrold J. Katz, Yuji Nishiyama, Scott Soames, Robert Steckerand, and Peter Tovey (1973). “Tacit Knowledge”. *The Journal of Philosophy*, 70(11), pp. 318–330.

Grice, H. Paul (1957). “Meaning”. *The Philosophical Review*, 66(3), pp. 377–388.

Harman, Gilbert (1973). *Thought*. Princeton University Press.

Haselanger, Pim, Andre de Groote, and Hans van Rappard (2003). “Representationalism vs. anti-representationalism: A debate for the sake of appearance”. *Philosophical Psychology*, 16(1), pp. 5–24.

Hauser, Marc D., Noam Chomsky, and W. Tecumseh Fitch (2002). “The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?” *Science*, 298, pp. 1569–1579.

Hawthorne, John (1990). “A note on ‘Languages and Language’”. *Australasian Journal of Philosophy*, 68(1), pp. 116–118.

——— (1993). “Meaning and Evidence: A Reply to Lewis”. *Australasian Journal of Philosophy*, 71(2), pp. 206–211.

——— (1994). “A Corrective to the Ramsey-Lewis Account of Theoretical Terms”. *Analysis*, 54(2), pp. 105–110.

Hill, Christopher S. (1991). *Sensations: A Defense of Type Materialism*. Cambridge University Press.

Horgan, Terence (1984). “Supervenience and Cosmic Hermeneutics”. *The Southern Journal of Philosophy*, 22(S1), pp. 19–38.

——— (1993). “From Supervenience to Superdupervenience: Meeting the Demands of a Material World”. *Mind*, 102(408), pp. 555–586.

Hornstein, Norbert (2009). *A Theory of Syntax: Minimal Operations and Universal Grammar*. Cambridge University Press.

Hornstein, Norbert and Paul Pietroski (2009). “Basic operations: Minimal Syntax-Semantics”. *Catalan Journal of Linguistics*, 8, pp. 113–139.

Janssen-Lauret, Frederique and Fraser MacBride (2020). “Lewis’s Global Descriptivism and Reference Magnetism”. *Australasian Journal of Philosophy*, 98(1), pp. 192–198.

Johnson, Kent (2014). “Methodology in generative linguistics: Scratching below the surface”. *Language Sciences*, 44, pp. 47–59.

Keiser, Jessica (2021). “On Meaning without Use”. *The Journal of Philosophy*, 118(1), pp. 5–27.

Kim, Jaegwon (1999). “Making Sense of Emergence”. *Philosophical Studies*, 95(1), pp. 3–36.

——— (2005). *Physicalism, or Something Near Enough*. Princeton University Press.

King, Jeffrey C. (2002). “Two Sorts of Claim about “Logical Form””. In: *Logical Form and Language*. Ed. by Gerhard Preyer and Georg Peter. Oxford University Press, pp. 118–131.

Knowles, Jonathan (2000). “Knowledge of grammar as a propositional attitude”. *Philosophical Psychology*, 13(3), pp. 325–353.

Kripke, Saul A. (1982). *Wittgenstein on Rules and Private Language*. Harvard University Press.

Langendoen, D. Terence and Paul M. Postal (1984). *The Vastness of Natural Language*. Basil Blackwell.

Laurence, Stephen (1996). “A Chomskian Alternative to Convention-Based Semantics”. *Mind*, 105(418), pp. 269–301.

Lewis, David K. (1969). *Convention: A Philosophical Study*. Harvard University Press.

——— (1970). “How to Define Theoretical Terms”. In *The Journal of Philosophy*, 67(13), pp. 427–446.

——— (1972). “Psychophysical and Theoretical Identifications”. In *Australasian Journal of Philosophy*, 50(3), pp. 249–258.

——— (1975). “Languages and Language”. In: *Language, Mind, and Knowledge*. Ed. by Keith Gunderson. Minnesota Studies in the Philosophy of Science, Vol. VII. University of Minnesota Press, pp. 3–35.
——— (1983). “New Work for A Theory of Universals”. *Australian Journal of Philosophy*, 61(4), pp. 343–377.
——— (1984). “Putnam’s paradox”. *Australian Journal of Philosophy*, 62(3), pp. 221–236.
——— (1992a). “Letter to John Hawthorne, 15 June 1992”. In *Philosophical Letters of David K. Lewis, Volume 2: Mind, Language, and Epistemology*. Oxford University Press, 2022.
——— (1992b). “Meaning without use: Reply to Hawthorne”. *Australasian Journal of Philosophy*, 70(1), pp. 106–110.
——— (1994). “Reduction of Mind”. In: *A Companion to the Philosophy of Mind*. Ed. by Samuel D. Guttenplan. Blackwell, pp. 412–431.
Loar, Brian (1976). “Two Theories of Meaning”. In: *Truth and Meaning: Essays in Semantics*. Ed. by Gareth Evans and John McDowell. Oxford University Press, pp. 138–161.
——— (1981). *Mind and Meaning*. Cambridge University Press.
Ludlow, Peter (2002). “LF and Natural Logic”. In: *Logical Form and Language*. Ed. by Gerhard Preyer and Georg Peter. Oxford University Press. pp. 132–168.
——— (2011). *The Philosophy of Generative Linguistics*. Oxford University Press.
Lycan, William G. (1981). “Form, Function, and Feel”. *The Journal of Philosophy*, 78(1), pp. 24–50.
——— (1987). *Consciousness*. The MIT Press.
——— (1996). *Consciousness and Experience*. The MIT Press.
Marr, David (1982). *Vision: A Computational Investigation Into the Human Representation and Processing of Visual Information*. W. H. Freeman.
Melnyk, Andrew (2003). *A Physicalist Manifesto: Thoroughly Modern Materialism*. Cambridge University Press.
——— (2016). “Grounding and the Formulation of Physicalism”. In: *Scientific Composition and Metaphysical Ground*. Ed. by Kenneth Aizawa and Carl Gillett. Palgrave Macmillan, pp. 249–269.
——— (2018). “In Defense of a Realization Formulation of Physicalism”. *Topoi*, 37, pp. 483–493.
Merrill, G. H. (1980). “The Model-Theoretic Argument Against Realism”. *Philosophy of Science*, 47(1), pp. 69–81.
Michaelis, Jens (2001). “Derivational Minimalism Is Mildly Context-Sensitive”. In: *Logical Aspects of Computational Linguistics*. Ed. by Michael Moortgat. Springer, pp. 179–198.
Myers, Scott (2000). “Boundary disputes: The Distinction between Phonetic and Phonological Sound Patterns”. In: *Phonological Knowledge: Conceptual and Empirical Issues*. Ed. by Noel Burton-Roberts, Philip Carr, and Gerard Docherty. Oxford University Press, pp. 245–272.
Neale, Stephen (1993). “Logical Form and LF”. In: *Noam Chomsky: Critical Assessments*. Ed. by C. Otero. Routledge, pp. 788–838.
——— (1994).“What is *Logical Form*?” In: *Logic and Philosophy of Science in Uppsala*. Ed. by Dag Prawitz and Dag Westerstahl. Springer, pp. 583–598.
Papineau, David (1996). “Theory-Dependent Terms”. *Philosophy of Science*, 63(1), pp. 1–20.
Peacocke, Christopher (1986). “Explanation in Computational Psychology: Language, Perception and Level 1.5”. *Mind and Language*, 1(2), pp. 101–123.
——— (1989). “When is a Grammar Psychologically Real?” In: *Reflections on Chomsky*. Ed. by Alexander George. Blackwell, pp. 111–130.
Pereplyotchik, David (2017). *Psychosyntax: The Nature of Grammar and its Place in the Mind*. Springer.
Piccinini, Gualtiero (2020). *Neurocognitive Mechanisms: Explaining Biological Cognition*. Oxford University Press.
Pietroski, Paul (2010). “Concepts, Meanings and Truth: First Nature, Second Nature and Hard Work”. *Mind and Language*, 25(3), pp. 247–278.
——— (2011). “Minimal Semantic Instructions”. In: *The Oxford Handbook of Linguistic Minimalism*. Ed. by Cedric Boeckx. Oxford University Press, pp. 472–498.
——— (2018). *Conjoining Meanings: Semantics Without Truth Values*. Oxford University Press.
Pietroski, Paul and Georges Rey (1995). “When Other Things Aren’t Equal: Saving *Ceteris Paribus* Laws from Vacuity”. *The British Journal for the Philosophy of Science*, 46(1), pp. 81–110.
Poeppel, David (2017). “The Influence of Chomsky on the Neuroscience of Language”. In: *The Cambridge Companion to Chomsky*. Ed. by James McGilvray. Cambridge University Press, pp. 155–174.
Polger, Thomas W. (2004). *Natural Minds*. The MIT Press.
Pollock, John L. (1989). *How to Build a Person: A Prolegomenon*. The MIT Press.

Putnam, Hilary (1960). “Minds and Machines”. In: *Dimensions of Mind: A Symposium*. Ed. by Sidney Hook. Collier Books, pp. 138–164.
Quine, Willard Van Orman (1960). *Word and Object*. The MIT Press.
——— (1972). “Methodological Reflection on Current Linguistic Theory”. In: *Semantics of Natural Language*. Ed. by Donald Davidson and Gilbert Harman. Springer, pp. 442–454.
Ramsey, Frank P. (1999). (1929).“Theories”. In: *Philosophical Papers: F. P. Ramsey*. Ed. by D. H. Mellor. Cambridge University Press, pp. 112–136.
Rattan, Gurpreet (2002). “Tacit knowledge of grammar: A reply to Knowles”. *Philosophical Psychology*, 15(2), pp. 135–154.
Ray, Greg (1995). “Thinking in L”. *Noûs*, 29(3), pp. 378–396.
Rescorla, Michael (2014). “A theory of computational implementation”. *Synthese*, 191(6), pp. 1277–1307.
Rey, Georges (1997). *Contemporary Philosophy of Mind: A Contentiously Classical Approach*. Blackwell.
——— (2003). “Chomsky, Intentionality, and a CRTT”. In: *Chomsky and His Critics*. Ed. by Louise M. Antony and Norbert Hornstein. Blackwell, pp. 103–139.
——— (2020). *Representation of Language: Philosophical Issues in a Chomskyan Linguistics*. Oxford University Press.
Scheer, Tobias (2010). *A Guide to Morphosyntax-Phonology Interface Theories: How Extra-Phonological Information is Treated in Phonology since Trubetzkoy's Grenzsignale*. De Gruyter.
Scheutz, Matthias (2001). “Computational Versus Causal Complexity”. *Minds and Machines*, 11, pp. 543–566.
Schiffer, Stephen (1993). “Actual-Language Relations”. *Philosophical Perspectives*, 7, pp. 231–258.
——— (2006). “Two Perspectives on Knowledge of Language”. *Philosophical Issues*, 16(1), pp. 275–287.
——— (2015). “Meaning and Formal Semantics in Generative Grammar”. *Erkenntnis*, 80(1), pp. 61–87.
——— (2017). “Intention and Convention in the Theory of Meaning”. In: *A Companion to the Philosophy of Language*. Ed. by Bob Hale, Crispin Wright, and Alexander Miller. Second Edition. Blackwell, pp. 49–72.
Schurz, Gerhard (2005). “Semantic Holism and (Non-)Compositionality in Scientific Theories”. In: *The Compositionality of Meaning and Content, Volume I: Foundational Issues*. Ed. by Edouard Machery Markus Werning and Gerhard Schurz. Ontos Verlag, pp. 271–284.
Schwarz, Wolfgang (2014). “Against Magnetism”. *Australasian Journal of Philosophy*, 92(1), pp. 17–36.
Shoemaker, Sydney (1981). “Some Varieties of Functionalism”. *Philosophical Topics*, 12(1), pp. 93–119.
Sider, Theodore (2011). *Writing the Book of the World*. Oxford University Press.
Stabler, Edward (1983). “How are grammars represented?” *Behavioral and Brain Sciences*, 6(3), pp. 391–421.
——— (2013). “Two Models of Minimalist, Incremental Syntactic Analysis”. *Topics in Cognitive Science*, 5(3), pp. 611–633.
Stalnaker, Robert C. (1984). *Inquiry*. Cambridge University Press.
Stich, Stephen P. (1971). “What Every Speaker Knows”. *The Philosophical Review*, 80(4), pp. 476–496.
——— (1978). “Beliefs and Subdoxastic States”. *Philosophy of Science*, 45(4), pp. 499–518.
Szábo, Zoltán Gendler (2012). “Against Logical Form”. In: *Donald Davidson on Truth, Meaning, and the Mental*, Ed. by Gerhard Preyer. Oxford University Press, pp. 105–126.
Tiehen, Justin (2016). “Physicalism Requires Functionalism: A new Foundation and Defense of the via Negativa”. *Philosophy and Phenomenological Research*, 93(1), pp. 3–24.
——— (2018). “Recent Work on Physicalism”. *Analysis*, 79(3), pp. 537–551.
Uriagereka, Juan (2012). *Spell-Out and the Minimalist Program*. Oxford University Press.
Waights Hickman, Natalia (2021). “(Implicit) Knowledge, Reasons, and Semantic Understanding”. *Mind and Language*, 36(5), pp. 707–728.
Weatherson, Brian (2013). “The role of naturalness in lewis's theory of meaning”. *Journal of the History of Analytical Philosophy*, 1(10), pp. 1–19.
Williams, J. R. G. (2008). “The Price of Inscrutability”. *Noûs*, 42(4), pp. 600–641.
——— (2015). “Lewis on Reference and Eligibility”. In: *A Companion to David Lewis*. Ed. by Barry Loewer and Jonathan Schaffer. Blackwell, pp. 367–381.
Wilson, Jessica M. (2005). “Supervenience-Based Formulations of Physicalism”. *Noûs*, 39(3), pp. 426–459.
——— (2014). “No Work for A Theory of Grounding”. *Inquiry*, 57(5–6), pp. 535–579.
——— (2018). “Grounding-Based Formulations of Physicalism”. *Topoi*, 37, pp. 495–512.
——— (2021). *Metaphysical Emergence*. Oxford University Press.

Yalcin, Seth (2014). “Semantics and Metasemantics in the Context of Generative Grammar”. In: *Metasemantics: New Essays on the Foundations of Meaning*. Ed. by Alexis Burgess and Brett Sherman. Oxford University Press, pp. 17–54.

How to cite this article: Balcarras, D. (2023). Functionalism and tacit knowledge of grammar. *Philosophical Perspectives*, 1–31. https://doi.org/10.1111/phpe.12179