Croatian Journal of Philosophy
Vol. XX, No. 60, 2020

Generative Linguistics Meets Normative Inferentialism: Part 1

DAVID PEREPLYOTCHIK*
Kent State University, Ohio, USA

This is the first installment of a two-part essay. Limitations of space prevented the publication of the full essay in present issue of the Journal. The second installment will appear in the next issue, 2021 (1). My overall goal is to outline a strategy for integrating generative linguistics with a broadly pragmatist approach to meaning and communication. Two immensely useful guides in this venture are Robert Brandom and Paul Pietroski. Squarely in the Chomskyan tradition, Pietroski's recent book, Conjoining Meanings, offers an approach to natural-language semantics that rejects foundational assumptions widely held amongst philosophers and linguists. In particular, he argues against extensionalism—the view that meanings are (or determine) truth and satisfaction conditions. Having arrived at the same conclusion by way of Brandom's deflationist account of truth and reference, I'll argue that both theorists have important contributions to make to a broader anti-extensionalist approach to language. What appears here as Part 1 of the essay is largely exegetical, laying out what I see as the core aspects of Brandom's normative inferentialism (§1) and Pietroski's naturalistic semantics (§2). In Part 2 (next issue), I argue that there are many convergences between these two theoretical frameworks and, contrary to first appearances, very few points of substantive disagreement between them. If the integration strategy that I propose is correct, then what appear to be sharply contrasting commitments are better seen as interrelated verbal differences that come down to different—but complementary—explanatory goals. The residual disputes are, however, stubborn. I end by discussing how to square Pietroski's commitment to predicativism with Brandom's argument that a predicativist language is in principle incapable of expressing ordinary conditionals.

Keywords: Generative linguistics, anti-extensionalism, normativity, inferentialism, predicativism, public language, communication.

* My deepest thanks to Daniel Harris, Ryan DeChant, Elmar Unnsteinsson, Eliot Michaelson, and Jessica Lynne Cooperrider for helpful advice on earlier drafts

Introduction

I take it that the correct approach to natural-language syntax is the one that Noam Chomsky outlined as early as the 1950s and, along with many others, has continually refined over the past seven decades. The ongoing research program of generative linguistics that his syntactic theorizing inspired has, in the fullness of time, yielded a diversity of impressive results. These include exciting and previously unimaginable empirical discoveries about the human capacity for language, both in broad scope—e.g., recursive generability and the principles-and-parameters model—and at the level of fine-structure (e.g., traces, parasitic gaps, etc.). However, as a theorist interested not only in syntax but also in semantics, I find myself in a difficult and somewhat awkward position.

Not to complain, but, you see, I happen to have learned my semantics from the work of Robert Brandom, and it's safe to say that I've drunk the Kool-Aid that he served up in his magnum opus, *Making It Explicit*. Having thus bought into *both* Chomsky's generative grammar *and* Brandom's normative inferentialism, I now find myself facing the daunting challenge of bridging the apparent chasm between the two. It may be that I'm utterly alone in this quandary, but in the course of the present discussion, I hope—perhaps somewhat perversely—to draw others into it as well.

There may never be an academic conference addressing the common themes and shared commitments of generative linguistics and normative inferentialism. For all that, the differences between them are, I believe, more boring—i.e., verbal or sociological—than is widely assumed. In what follows, I'll argue that, contrary to first appearances, there are actually very *few* points of substantive disagreement between them. The residual differences are stubborn, to be sure, but this can only be appreciated after a suitably wide collection of background agreements is put into place. I devote the first half of the present discussion to this latter task.

Squarely in the Chomskyan tradition, Paul Pietroski's recent book, *Conjoining Meanings*, offers an approach to natural-language semantics that rejects foundational assumptions widely held amongst philosophers and linguists. In particular, he argues against the view that

of this paper; thanks also to Nate Charlow, Matt Moss, and other participants of the 2018 Ontario Meaning Workshop, for stimulating exchanges on related topics. That said, if some of what I claim here is false, then these people are all partly to blame. Ditto for cases where I say something awkwarder than usual; they should of caught that. My sincere gratitude, lastly, to all of the participants at the two symposia at which I was fortunate to present earlier versions of these ideas, including the 2018 Croatian Meaning Workshop. From the latter group, I owe special thanks to Paul Pietroski (obviously), Michael Devitt, John Collins, Anna Drożdżowicz, and, most of all, to Dunja Jutronić—not only for inviting me to participate in the workshop, but for her saintly and near-inexhaustible patience with me throughout this project. Many thanks also to Nenad Miščević for allowing the second half of the present essay to appear in the next issue of the Journal.

meanings are, or even determine, extensions. The latter include such familiar semantic properties as truth conditions, satisfaction conditions, and denotation/reference. Having arrived at the same conclusion by way of Brandom's deflationist account of truth (§1.3), I began to glimpse the possibility of a fruitful merger of ideas. The present essay represents a first pass at integrating the generative linguist's empirical insights about human psychology with the broadly pragmatist framework about mind and language. I'll argue that both have important contributions to make to our overall understanding of language. The easy part is spelling out what; the harder part is assessing the residual disagreements.

Here's the overall plan:

In §1, I survey a range of core commitments that jointly constitute Brandom's philosophical project—most centrally, his normative pragmatics, inferentialist semantics, and substitutional syntax. Along the way, I note his intellectual debts to David Lewis, including the large-scale explanatory goals that animate Brandom's inquiry. These, I later argue, are in many ways orthogonal to Pietroski's concerns. The latter claim figures in my broader argument that the two approaches can be fruitfully combined, potential protests from both sides notwithstanding.

In §2, I outline Pietroski's position, focusing on his explanatory aims, his empirical methodology, and the substance of his proposal for a theory of human semantic competence. I take up Pietroski's arguments against Lewis's approach to natural language, with the aim of showing that Brandom's theoretical goals differ sufficiently from Lewis's to inoculate him against Pietroski's criticisms. Turning to Pietroski's discussions of Frege, I point out that his work in cognitive science undercuts some of Brandom's claims in "Why Philosophy Has Failed Cognitive Science". As we'll see, far from ignoring Frege, Pietroski incorporates many of his insights into an empirical account of psychological processes. Nevertheless, some of the Fregean lessons that Brandom emphasizes do *not* seem to have moved Pietroski. In later sections, I explore some possible reasons for this.

This will conclude Part 1 of the essay. Limitations of space prevented the publication of the full essay in present issue of the Journal. The second installment will appear in the next issue, 2021 (1). Here is a preview of what it contains:

I devote §3 to a survey of the core commitments that Pietroski and Brandom have in common. As already noted, both reject truth-conditional semantics and seek to develop alternative frameworks for semantic theorizing. Similarly, I'll point out, the alternatives they propose can both be seen as taking referential *purport* to be the distinctive feature of language. This contrasts with received views in (meta) semantics that focus, in the first instance, on referential *success*, leaving reference *failure* (e.g., empty names) for special treatment, as a blessedly rare "defective" case. This is related, I suspect, to the further convergence between the Brandom and the Pietroski on the proper

treatment of the distinction between *de dicto* and *de re* constructions. Last, but no less important, is their common rejection of the idea that communication requires an *identity* between the meanings expressed by a speaker and those grasped by the hearer. I'll note that this shared commitment undercuts some of the main arguments against meaning holism—an inherent feature of Brandom's inferentialist account.

A discussion of the differences between Pietroski and Brandom occupies the remainder of the essay. At first glance, the contrast between them seems as vivid as any in the field. Brandom's project is explicitly *normative*, pursued largely from the armchair, and aims to provide an account of concept users *as such*—not just humans, but *any* conceivable concept-mongering creature (or artificial system). By contrast, Pietroski's project is avowedly *descriptive*, constrained by *empirical* data, and aims to provide an account of *actual humans*—particularly, the "Slangs" that they naturally produce and consume. (Pietroski uses 'Slang' as a catch-all term for the natural languages that children acquire.) These differences ramify quickly. For instance, Brandom's focus on social practices of communication seems to be at odds with Pietroski's "individualistic" methodology, which is characteristic of generativist views more broadly (Chomsky 2000; Collins 2012). This contrast seems particularly sharp in light of Brandom's commitment to the existence of (something like) public languages—at least in the sense of productive and flexible norms governing the communal practice of "giving and asking for reasons" (GOGAR). All of these issues are discussed in §4, with the aim of gradually blunting the force of what initially appear to be quite sharp differences.

If one insists on seeing Brandom's and Pietroski's inquiries as targeting a common subject matter, then one is sure to view the differences between them as substantive theoretical *disagreements*. They do, after all, use the same term, 'language', which at least *suggests* that they're talking about one and the same phenomenon. But this way of viewing the situation is optional, at best. We've learned by now to stop assuming that theorists are targeting the same phenomenon simply on account of their using homophonous terms. Two theorists can press the same bit of folk vocabulary, e.g., 'meaning' or 'concept', into more weighty theoretical labor in quite different ways. That being so, one can just as well see the inferentialist and the generativist as addressing *different* (though undoubtedly related) topics—each providing insights about his chosen domain of inquiry, and leaving the rest of us to wonder how those insights might be integrated, or at least brought to bear on one another. This is the strategy I'll recommend throughout the present discussion.

In arguing that the two are, at the very least, logically *compatible*, one owes an account—or, at a minimum, a broad *sketch* of an account—of the theoretical *relation* between them. The proposal that I'll develop is that Brandom's explanatory ambitions differ from Pietroski's in precisely the ways that are paradigmatic of "*inter-level* theoretical

relations". My suggestion is that Brandom is attempting to furnish a *high-level description* of a quite general phenomenon—*language use, as such*. This presupposes that lower-level implementations of the relevant generalizations can vary widely.[1] Pietroski's theoretical aims are, though certainly more exciting from an empirical standpoint, a good bit narrower than Brandom's, in that they deal exclusively with the human case. As with any lower-level account of a more general phenomenon, Pietroski's view is compatible, in principle with any number of higher-level descriptions of language. Thus, while it's incorrect—or, at any rate, misleading—to say that the two accounts are strictly orthogonal to one another, the fact is that each places very few constraints on the other. It *should* be surprising, then, when we find substantive points of contact between them, whether these be points of contention or convergence.

Still, even if my suggestion is right that the two frameworks are more compatible than might initially appear, we must face up to the residual differences that credibly threaten my reconciling project. The most stubborn of these, which I accordingly leave to the very end, has to do with Pietroski's arguments for a specific version of *predicativism*—roughly, the view that all subsentential entities are best treated as predicates, not (say) singular terms (§2.4). To be sure, Pietroski's commitment to this view is not a central aspect of the overall generativist enterprise. Rather, it's a tendentious empirical hypothesis, for which he offers correspondingly forceful arguments. That being so, if it were to turn out that the hypothesis is false, generative linguistics would go on without a hitch. Still, I focus on this issue because it raises much larger questions about how to treat subsentential entities, not just at the level of semantics, but also at the level of *syntax*.

The notion of syntax that Brandom employs is *substitutional*—not in the sense of "substitutional quantification" (though he endorses that too, on independent grounds), but, rather, in the sense that he takes *sentences,* i.e., expressions of *full thoughts*, to be the primary vehicles of meaning. On his view, subsentential items (words, morphemes, etc.) are the products of taking a "substitutional scalpel" to the antecedently-interpreted *sentential* unities. We'll take a first stab at unpacking the scalpel metaphor in §1.5, and then come back to it in more detail in §4.3. While this substitutional approach *may* have a home in a Fregean semantics—Pietroski's powerful arguments notwithstanding—there is no *obvious* sense in which it can be legitimately applied to the syntax of human languages.

The trick that I pull repeatedly throughout the second half of this essay—namely, that of relegating the two inquiries to distinct theoretical "levels"—doesn't get much of a grip here. For, it's a requirement of any such picture that an account pitched at the (relatively) higher level

[1] Indeed, as Fodor (1975) points out, without substantive constraints from the higher-level account, such "realization bases" might differ *indefinitely*.

of analysis should be *compatible*, at least in *core* respects, with any lower-level account of the "realization base". But Brandom's substitutional syntax seems, even upon close scrutiny, to be not only *different* from generative grammar, but decidedly *at odds* with it. I strongly suspect that this different comes down to a methodological conflict between the generativist's "bottom-up" treatment of subsentential items and the inferentialist's "top-down" alternative.[2] In §4, I discuss and evaluate some ways of viewing this disagreement, arriving ultimately at the following bittersweet conclusions.

Despite strong grounds for optimism about the possibility of integrating normative inferentialism with the up-and-running research program of generative linguistics, it must be admitted that rendering Brandom's substitutional approach to syntax compatible with the going theories in generative grammar (e.g., Chomsky's minimalist program) presents an obstinate challenge. Though I can think of no reason that the challenge is insuperable *in principle*, it's nevertheless the case that I am not, at present, equipped to meet it myself. Perhaps others can do better in this regard—a task I invite, encourage, and exhort philosophers of language to undertake, on the strength of the positive arguments adduced here.

1. *Brandom: From normative pragmatics to inferentialist semantics (and back)*

Introduction

Robert Brandom's philosophical project is grand in both scope and ambition.[3] The resulting theoretical framework has a number of moving parts, to put it mildly. In this section, I'll lay out what I take to be the central commitments of his "normative inferentialism", with particular focus on those that pertain to the broader goals of this discussion—i.e., the proposal to integrate the generativist and inferentialist research programs.

Although this section is intended to be largely exegetical, postponing critical evaluation to §4, I should emphasize that it wouldn't matter to me very much if Brandom wouldn't put things *quite* the way that I do below. While his account is by far the most well-worked out version of normative inferentialism, and hence immensely useful as a guide

[2] See Collins (2012) for a detailed and rewarding discussion of this issue. The contrast that Collins draws between sentence-first and word-first approaches (my terms, not his) serve, if anything, to *sharpen* the contrast that I'm worried about here. My hunch is that coming to grips with Collins' conclusions in that work will be crucial to resolving the issues I raise in §4.3.

[3] The picture I sketch in what follows is drawn mostly from the material in Brandom (1994, 2000, and 2008). In what follows, I'll occasionally abbreviate these to *MIE* (*Making it Explicit*), *AR (Articulating Reasons)*, and *BSD (Between Saying and Doing* 2008*)*, respectively. The final sections of this essay also make heavy use of the material in "Why Philosophy Has Failed Cognitive Science" (in Brandom 2009).

in this area of inquiry, my intent is not so much to capture every nook and cranny of one particular theorist's gargantuan philosophical system. Rather, the goal of this section is to present and motivate an account of language that is attractive enough to warrant comparison with Pietroski's independently attractive proposals about natural-language semantics (§2). Ultimately, it's only by comparing them that we can put ourselves in a position to clearly assess the merits of either, let alone to contemplate their integration.

1.1. *Methodology and explanatory aims*

The central questions, for Brandom, are about what *constitutes* language use—not by humans, necessarily, but by any natural creature or artificial system to which attributions of a linguistic competence are warranted. Given that normal adult humans have mastery of at least one natural language, Brandom's account will, in *some* sense, apply to us as well—though perhaps only as a particular instance of a much more general phenomenon. Still, the inquiry he undertakes is not a straightforwardly psychological one; nor is it pitched as a historical or anthropological hypothesis. Rather, the idea is that, armed with a philosophically sophisticated and conceptually articulated account of "what the trick *is*"—where "the trick" is language use, broadly construed—an empirical scientist (a linguist, psychologist, or artificial intelligence researcher) can ask more detailed questions about "how the trick *happens to be done*" by one or another creature or artifact.

This latter kind of research is bound to yield an increasingly refined picture of some particular *type* of linguistic competence—paradigmatically, the human type (though one often hears of impressive progress in interpreting the languages of other social creatures, e.g., prairie dogs and dolphins). Thus, although the account is intended to apply far beyond the human case—to aliens, robots, and other terrestrial animals—there is no commitment on Brandom's part to the effect that empirical findings can have no bearing *whatsoever* on his philosophical claims. Nor does he hold that legitimate empirical inquiry can take place only *after* a credible philosophical account has been supplied. This is just one sense in which normative inferentialism and empirical science, including generative linguistics, are not in competition.

One might wonder, at this point, how far removed Brandom's project is from empirical considerations. Perhaps *too* far? True, many of the thinkers with whom his work most directly engages did *not* conceive of their inquiries on the model of empirical theorizing. In fact, the three philosophers who figure most centrally in *MIE*—namely, Kant, Frege, and Wittgenstein—all famously made a point of distancing themselves from natural science. But we have to keep in mind, here as elsewhere, that Brandom's theoretical framework is "a house with many rooms". Thus, we need not take a one-size-fits-all approach to this question; indeed, there are good reasons not to.

In providing detailed analyses of linguistic constructions in distinctively human languages—*de dicto/re* reports, truth-talk, (in)definite descriptions, indexicals, deixis, anaphora, modals, quantification, predicates, and singular terms—Brandom relies on exactly the same stock of empirical considerations that one finds in standard semantics texts (e.g., Heim and Kratzer 1998). Such data are often cast in terms of intuitive judgments concerning the *truth*-values or *truth*-conditions of a target sentence, often in the context of auxiliary assumptions that are supplied by an accompanying vignette. But these very same data can equally well be recast as competent speakers' judgements concerning (not truth but) *inferential proprieties*—e.g., what one is required or permitted to infer on the basis of the assumptions supplied (in combination with all prior background information that is assumed to be common-knowledge).[4]

A similar treatment can be given of metalinguistic intutions/judgments concerning minimal pairs and sentential ambiguity. Two sentences differ in meaning, on the inferentialist account, just in case they differ in respect of the inferential proprieties that govern their use. Likewise, a sentence is ambiguous just in case it is capable of playing two distinct/incompatible inferential roles. And, though he doesn't, to my knowledge, ever discuss the phenomenon of polysemy, I presume that Brandom would treat is as a case of *overlapping* inferential properties—perhaps ones that meet some further normative or inferential conditions.

The fact that Brandom often appeals to precisely such data in developing his inferentialist semantics suggests that at least *this* aspect of his view is firmly grounded in empirical fact. But, as I'll emphasize later, Brandom's use of these data is not empirical but, rather, *illustrative*. He is, in other words, using examples from English—indeed, almost exclusively English, in contrast with the generativist's cross-linguistic methodology—as case studies that, according to him, exemplify more general phenomena. Thus, the English-language examples that he occasionally provides—a bit too rarely, one might lament—serve not as empirical data for a scientific or naturalistic hypothesis. Rather, the most charitable reading of his appeals to such examples casts them as attempts to help us to get a conceptual grip on features of language(s) *as such*.

Still, it must be admitted that other aspects of Brandom's overall picture are far less tethered to the facts on the ground. Presently, we'll see that his normative pragmatics—to which his downstream proposals, including inferentialism, are conceptually subordinate—makes only very minimal empirical assumptions. For instance, it takes for granted that complex social creatures came into being *somehow or other*—e.g., via evolution by natural selection, or by deliberate engi-

[4] My thanks to Eliot Michaelson and Daniel Harris for helping me to see the points in this paragraph more clearly.

neering, as with AI. Brandom likewise presumes that the behavioral control systems of such creatures/artifacts—brains, motherboards, or what have you—work *somehow or other*; else they couldn't perform the behaviors that institute social norms, given reasonable naturalistic constraints. Still, putting aside such near-vacuous assumptions, this aspect of Brandom's project can fairly be described as an armchair enterprise.[5] Whether that's something to hold against it—e.g., on the basis of one or another naturalist scruple—is something we can only ascertain only after we've surveyed the details of his overall proposal. To these we now turn.

1.2. *Normative pragmatics*

Brandom begins by situating all linguistic practices in the wider realm of activities that are, in some sense, *rule-governed*. The notion of a rule plainly stands in need of careful articulation. Given our philosophical history, two options immediately suggest themselves—what Brandom calls "regulism" and "regularism". According to the *regulist*, rule-following is a matter of obeying rules that one can *explicitly* formulate or comprehend. By contrast, the *regularist* holds that rule-following is a matter of being disposed to behave in a way that accords with one or another empirical *regularity*. Brandom rejects both of these alternatives, though, as we'll see, his own account is an attempt to split the difference between them.

Regulism is vitiated, he argues, by the fact that obeying an explicit rule—e.g., a dictionary definition or an academic/prescriptive convention of grammar or style—requires first *interpreting* the rule. This, in turn, requires *deploying concepts*—a version of precisely the phenomenon that we're seeking to explain. The regularist option, he contends, faces a distinct challenge. A pattern of behavior—whether finite or infinite, actual or potential—can be either a *successful* or a *defective* case of following a given rule. So, with regard to *any* pattern of performance, the question always stands: has the rule in question been followed *correctly*? Assessments of correctness are, on Brandom's view, inherently normative. Yet, what the regularist offers is a purely descriptive account, couched in the language of cognitive or behavioral dispositions (and perhaps other, related alethic modal notions; see §3.1).

Brandom's positive proposal insists on the normativity of assessment, characteristic of the regulist view, but jettisons the requirement that the rules in question be manifested explicitly from the very outset. Though rules of practice can *eventually* come to be articulated—i.e., "made explicit"—by a community of concept-using creatures, such rules are, in the first instance, *implicit* in the social practices of the creatures in question.

[5] Notably, a former student of Brandom's, John MacFarlane, has programmed a version of "the game of giving and asking for reasons" (GOGAR) for the popular online game, *The Sims*. https://www.johnmacfarlane.net/gogar.html

Why social? Why not, instead, the practices of an individual? Put crudely, the reason is that isolated individuals cannot, in principle, serve as a normative check on their own judgment and behavior. In order for a creature to be so much as *subject* to normative assessment, its behavior must take place in a context where other creatures respond to it in ways that signal social (dis)approval. How numerous and long-lasting must the social relations be in order to institute a social practice, properly so called? Brandom's answer, which will become relevant later in the discussion (§4.1), may be somewhat surprising. Terms like 'social' and 'communal' bring to mind a relatively large group of creatures. But, when he presses these folk terms into theoretical service, Brandom's official view is far less committal than that. All it really requires is a *dyadic* "I-thou" relation—i.e., a case of mutual recognition, in respect of authority and responsibility, on the part of at least *two* creatures/systems. Such relations of mutual recognition can be merely *implicit* in the 2+ creatures' overt practices toward one another—e.g., one or another type of *social sanction*.

In the most basic case, social sanctions come down to either naked violence or the provision of necessities—beatings and feedings. But once a social practice becomes sufficiently complex, it comes to include not only such "external" sanctions, but also "internal" ones—e.g., initially, the granting of privileges and later the exchange of *tokens* of privilege. (We are invited here to imagine a special fruit that permits its bearer to enter a particular territory without being attacked by the creatures who guard it.) With each additional layer of interwoven internal sanctions, the community becomes increasingly ripe for instituting not merely norms of practical action, broadly construed, but the more specifically *linguistic* norms of assertion and demand. I include the latter here, not because Brandom ever treats it in detail, but on account of his frequent invocation of the trope "the game of giving *and asking for* reasons" (my emphasis, to be sure). While the "asking" part seems to deserve equal attention, Brandom takes the norms governing *assertion* to be the "downtown" of language—a backhanded adaptation of Wittgenstein's metaphor of language as a city with *no* downtown.

Having argued that assertion is the fundamental pragmatic notion, Brandom goes on to give an account of it in terms of the normative social statuses of commitment and entitlement. To illustrate these notions, let's work through a hypothetical example of how the game of giving and asking for reasons might be played amongst a group of primitive hominids—or, for that matter, current prairie dogs.[6]

Suppose a creature produces a public token in a social context, and that this act has—if only in that context—the pragmatic significance of explicitly committing the creature, in the eyes of its community, to its

[6] I am acutely aware that, despite my efforts at rendering the scenery in a plausibly naturalistic light, the example is not only fictional but transparently artificial in countless respects. I trust these won't matter for the sake of making the key points accessible.

being the case that the enemy is approaching. Each of the other creatures in the group evaluates this commitment, assessing it as correct or incorrect on the basis of their own commitments, explicit or otherwise. In doing so, they take a *normative stance* toward the "speaker", whom the group might then treat as being *entitled* to that commitment. The speaker can be entitled to a claim either by default—e.g., in cases of joint perception or contextually-relevant common knowledge—or by having undertaken prior commitments that jointly warrant the one in question. Initially, such normative attitudes are implicit in the assessors' overt treatment of the speaker. The group might, for instance, shift its attention in the direction of the speaker's gaze upon hearing 'Enemy!". If the enemy is indeed approaching in a way that is perceptually evident to the group, the entitlement is thereby secured.

Suppose now that the same speaker, now quite anxious, produces another token—e.g., "Run!" This counts not only as an explicit commitment to a (potentially joint) plan of action, but also an implicit commitment to the goodness of the inference from "Enemy!" to "Run!" And while the members of the group assessed the speaker as being entitled to the first claim, they may well go on to treat this new explicit commitment, i.e., "Run!", as patently unwarranted in the circumstances. Again, in the most basic cases, this "treatment" or "stance" toward the speaker can take the form of overt actions on the part of other group members—e.g., grabbing hold of the speaker and keeping them in place. Another form of response might be the production of overt tokens that commit the group, including the original now-frightened speaker, to an incompatible plan—e.g., "Stay!" and "Fight!". This normative incompatibility is itself implicit in the overall communal practices of the group.

From these primitive beginnings, Brandom suggests, a practice can evolve in such a way as to allow for speakers to make *explicit* their commitments regarding the goodness—or, in his terms, "material propriety"—of the inferences that were previously only *implicit* in their practices. Paradigmatically, this is achieved by introducing an expression that has the significance of a conditional. For instance, we can imagine a newly-evolved creature—call it v2.0—that has achieved what Brandom calls "semantic self-consciousness". This involves discursively representing not just enemies and escape strategies, but also *inferential relations between claims*. Creature v2.0 can make explicit its commitment to the goodness of a particular inference by producing a token that has the significance of the conditional, e.g., "If Enemy then Run!"[7] Further elaborations of v2.0's language are manifested with the introduction of new bits of logical vocabulary, all of which serve to express commitments regarding various inferential relations. For instance, negation expresses the relation of inferential *incompatibility* between claims (see below). This is the core thesis of what Brandom calls "logical expressivism".

[7] Note that the force operator, '!', is stripped off from the atomic propositions, "Enemy!" and "Run!", when the latter are embedded in a conditional. This point comes to the foreground in §4.2.

Turning to more complex logico-semantic devices, consider the phenomenon of indirect discourse. Brandom proposes that *de dicto* reports of "what was said" are used to make explicit the *attribution* of commitments to oneself or to others. For instance, my asserting "Dan said that p" makes explicit *my* commitment to Dan's having explicitly undertaken *his own* commitment to the effect that *p*. In a similar fashion, epistemic vocabulary can play the role of making explicit a speaker's *assessment* of someone's *entitlement* to the commitments that they've undertaken. For instance, were I to upgrade my assertion to "Dan *knows* that *p*," I would not only be *attributing* the commitment to Dan, but also committing *myself* to it, and explicitly representing him as being entitled to it. Brandom (2001: ch. 4) points out that these three *pragmatic* aspects of knowledge *attributions*—entitlement-attribution, commitment-undertaking, and commitment-attribution—correspond, in that order, to the three elements of the traditional Justified-True-Belief account of knowledge.

What about deontic modal terms? Brandom argues that these serve to make explicit one's commitment to a plan of action. In saying "I ought to φ", I make explicit my commitment to a plan to φ. Similarly, terms like 'must' can be used to make explicit an inferential propriety that is insensitive to changes in auxiliary assumptions, up to some boundary condition—e.g., natural or legislated laws. Thus, assertions like "In order to light the wick safely, one must first clean one's hands" serve to make explicit a speaker's commitment to the material propriety of the inference from "The wick-lighting is safe" to "Your hands are clean," irrespective of what commitments they have concerning a wide range of possible auxiliary assumptions, such as "It's raining elsewhere," "I never met my grandfather," and "Child-trafficking is a serious problem." Whether or not these latter are included in one's set of commitments, the inference from "The wick-lighting is safe" to "Your hands are clean" is ostensibly good. Of course, auxiliary commitments that conflict with one's views on *natural laws*—e.g., "Gasoline burns when lit"—would render the inference materially invalid. That's what talk of "boundary conditions" is intended to capture. In the case of the purely nomic reading of 'must', as in "Oxygen must be present for combustion," the inference from "Combustion is occurring" to "Oxygen was present" is good under substitution of virtually any commitment, other than those regarding physical or chemical law.

In this vein, Brandom provides rich pragmatic analyses of a variety of other "vocabularies", including the (meta)semantic devices 'represents' and 'is about', as well as indexicals, alethic modals, anaphoric pronouns, and *de re* attitude reports. Some of the details of these analyses will emerge throughout the present discussion, and I'll devote special attention to his account of *de re* constructions in §3.4. For the moment, the case that's most important to examine is that of 'true', as this bears directly on Brandom's rejection of truth-conditional semantics—arguably the main negative contention that he shares with Pietroski.

Brandom offers a clear and well-motivated alternative to the standard treatment of truth in terms of "correspondence"—a notoriously vexed notion that lay at the heart of truth-conditional semantics. Instead he develops a refined version of the "deflationary" approach of truth and reference—arguably the best on the market, as we'll see presently.

1.3. *Deflationism about truth and reference*

In asserting that things are thus-and-so, a speaker takes on the pragmatic normative status of a discursive commitment. What is it, then, for *another* person to say that the first speaker's assertion is *true*? The normative pragmatic answer is this: in asserting that some claim is true, one not only *ascribes* a commitment to the speaker who made it, but one also *undertakes that commitment oneself*.[8] This allows for the possibility—enormously useful in social practice—for a speaker to take on commitments that they cannot at present articulate. The inability may be either to memory loss ("I don't remember exactly what she said, but it was definitely true") or to time constraints ("She gave a long speech; I'm not going to repeat the whole thing now, but everything she said was obviously true"). In the most interesting cases, complete articulation is impossible within physical limits, the set of commitments being literally infinite, as with "The theorems of Peano arithmetic are all true".

Thus, on Brandom's view, the term 'true' and its cognates ('truth', 'correctness', etc.) all serve a distinctive *expressive* function, without which branches of discourse such as mathematics would be impossible. Specifically, these terms all serve to express a commitment to something already asserted (or, at any rate, assert*ible*). Brandom thus labels this an "expressivist" account of truth, of a piece with his more general expressivist approach to logical vocabulary. The semantics for truth ascriptions is elaborated still further in light of his discussion of anaphora (§1.7). The notion of inter-sentential (and inter-speaker) anaphora will allow us to appreciate how truth ascriptions can have the pragmatic function of allowing the *inheritance* of commitments and entitlements across inter-personal exchanges.

If this pragmatic expressivist account of our use of 'true' (and related terms) is correct, then there is no obvious reason why anything further needs to be said about *truth*. The latter is often conceived of as a metaphysical language-world relation—the very one denoted/satisfied by the term 'truth' and its cognates. But there seems to be no explanatory work for which such a relation is *obviously* indispensable—neither in semantics nor, Brandom argues, in any other area of theorizing. This puts his view in the same camp as other versions of "deflationism" about truth—particularly, the well-known disquotational (Quine

[8] In the special case where the speaker and the assessor are identical—as in, "What I'm saying is *true*!"—the commitment is both redundant and guaranteed, though the term is still useful, in such cases, if only for emphasis and the like.

1970), minimalist (Horwich 1990), and prosentential accounts (Grover, Camp, and Belnap 1975). I take all of these to share the following core commitments:

(i) a rejection of any account/analysis of truth in terms of correspondence, coherence, or warranted assertibility, on the grounds that truth is *not a relation* (of *any* kind)

(ii) the aim of casting the notion of meaning as *explanatorily prior* to that of truth, both in semantics and elsewhere (including metaphysics, epistemology, and ethics)

The differences between disquotationalism, minimalism, and prosententialism have mostly to do with matters of detail, such as whether to ascribe truth to sentences or to propositions, or how exactly to interpret Tarski biconditionals, liar sentences, and quantified truth ascriptions. These disputes are all strictly irrelevant for our purposes. What's important here is that Brandom's version of deflationism is designed to claim the virtues of each of these prior accounts, without succumbing to the technical objections that have been lodged against them. The three main improvements he suggests are (i) subordinating the semantics of truth ascriptions to his brand of normative pragmatics, (ii) paying closer attention to the *syntax* of truth ascriptions, especially their intersentential anaphoric structure (§1.7), and (iii) extending the deflationist account to other semantic notions, including reference, satisfaction, and *de re* representation. While (i) is a straightforward application of Brandom's broader strategy, and (ii) serves largely to immunize his version of deflationism from extant objections, (iii) strikes me as a genuine extension of Brandom's normative pragmatics, allowing it to handle both sentential *and sub*sentential expressions.

The notions of truth and reference are plainly central to the project of truth-conditional semantics. Thus, many have noted that a deflationist account of these notions requires a radical re-thinking of what shape a formal semantic theory should take. In this regard, we now have an embarrassment of riches. In addition to old-school proposals about warranted assertibility, and the pragmatists' short-lived "success semantics" (see Brandom 2009 for critique), we now have the benefit of more modern proposals, including both Pietroski's cognitivist account (§2) and Brandom's inferentialism. Let's examine the latter.

1.4. *Inferentialist semantics*

Having situated assertional practices within the broader sphere of rule-governed social activity, Brandom has introduced his key *pragmatic* notions of commitment and entitlement. He goes on to show how these normative statuses, taken together, can be used to construct a *semantic* theory, whose business it is to explain (in some sense) *how* linguistic expressions can come to play the roles that they do in a community's assertional practices.

In familiar fashion, the explanation goes by way of assigning "meanings" or "semantic values" to expression types. But, in keeping with his other commitments, Brandom does *not* equate meanings with truth conditions, sets of possible worlds, pragmatic success conditions, or assertibility criteria. Rather, he subordinates his semantic theory to the normative pragmatics just outlined, by treating meanings as the *inferential proprieties* that govern the use of linguistic expressions. Slurring over a considerable mass of detail, we can summarize the proposal as follows:

Inferentialism: For a given propositional expression, 'P', the meaning of 'P' can be modeled as the set of sets of other propositional expressions that

(i) entitle one to 'P' in the presence of (various sets of) auxiliary commitments,

(ii) commit one to 'P' in the presence of (various sets of) auxiliary commitments, as well as those to which

(iii) 'P' commits one in the presence of (various sets of) auxiliary commitments, and

(iv) 'P' entitles one in the presence of (various sets of) auxiliary commitments.

A particularly useful *compound* inferential relation turns out to be that of *incompatibility*, wherein taking on one commitment precludes a speaker from becoming entitled to another. In this sense, a commitment to "Herbie is a dog" is incompatible with entitlement to "Herbie is a bird". Brandom (2008: ch. 5) shows how to build a *modal* propositional semantics on the basis of just this incompatibility relation, treating the negation of a claim, for instance, as the minimal set of commitments that are incompatible with it. Here again, the details are illuminating, but only one significant upshot bears highlighting for present purposes.

Casting the meaning of an expression in terms of its inferential proprieties vis-à-vis *other* expressions plainly commits one to meaning holism. A common charge against theories of a holist stripe is that they founder on the rock of compositionality. For instance, Fodor and Lepore (1992) famously argue that inferential roles don't compose, whereas meanings do; *a fortiori*, meanings can't be inferential roles. But the formal incompatibility semantics developed by Brandom (2008) provides a direct counterexample the main premise of this argument, by demonstrating how a inferentialist semantics *can* in fact *provably* meet reasonable compositionality constraints, at least in the modal propositional case.[9] In any event, we will see that there are other reasons to reject Fodor and Lepore's arguments.

[9] Although, as of this writing, an analogous proof for the *quantificational* case remains elusive, I am aware of no principled reasons for thinking that such a proof won't emerge—if not tomorrow, then *someday*. As will become clear throughout, I adopt a resolutely optimistic attitude toward such matters.

1.5. *Substitutional "syntax"*

We've now put on the table both a normative pragmatics and an inferentialist semantics. However, it's relatively uncontroversial that only *proposition*-sized expressions can enter *directly* into inferential relations as premises and conclusions.[10] That being so, we still need to say how *sub*sentential bits of language can have meanings of their own. Identifying subsentential expressions will allow us to explain how such expressions can go on to contribute to the indefinitely many assertions that a creature like us can interpret and produce. While there is no *conceptual* barrier, on Brandom's picture, to a community of creatures/robots using a language with only *finitely* many complex expressions, our own case plainly illustrates that languages can and do come in varieties that admit of *productive* generation. So, while a first-pass presentation of the inferentialist approach is best conducted in terms of a community of creatures that uses a finite language—such as might easily be found in (extra)terrestrial nature or constructed in a robotics laboratory (e.g., AIBO dogs)—it does *not* follow, and is not true, that the inferentialist program abdicates the responsibility of explaining the productive nature of *some* languages. Quite the contrary; Brandom takes his account of subsentential meanings to constitute one of the core achievements of the inferentialist program.

The primary notion of an inferentialist semantics for subsentential expressions is that of *substitution*, which Brandom inherits from (a reconstructed time-slice of) Frege. Starting with a finite stock of sentence types, Σ, each of whose free-standing (i.e., unembedded) uses have the default pragmatic significance of performing an assertion, we can ask whether any members of Σ can be treated as *substitutional variants* of any others. Keeping to the level of naïve intuition, the sentence 'David admires Herbie' is a substitutional variant of 'Jessica admires Herbie'. We'll see more about how this works in a moment, but the key takeaway point is this: if a sentence has a set of substitutional variants, then we can, *to that extent*, discern its subsentential structure. That is, by relating one sentence to another inferentially via substitution, we can notice and distinguish re-combinable subsentential expressions *within* the sentences of the language. Let's work through an example.

Take the sentence 'David admires Herbie' and chop it up any way you like, in respect of phonology, orthography, or whatever surface-level features happen to be relevant to the language at hand.[11] One way of doing so will yield 'Herbie' as a proper part; another yields '…

[10] For a dissenting view, see Stainton (2006).

[11] We'll do things in terms of orthography here (given the medium), but phonology is plainly the more primitive of the two in the human case, both phylo- and ontogenetically, as textbooks in empirical linguistics have long emphasized. For future robots, the medium will likely be something else—perhaps some descendent of TCP/IP. This would require adapting the substitutional techniques to that particular case.

…'erbi'...'. Now do the same with every other member of Σ, where the latter is assumed to be finite.[12] This yields a set of subsentential bits, Γ, consisting largely of nonsense like …'dmire'… and …'vid admi'…. With this in hand, go back to 'David admires Herbie' and substitute any other member of Γ (or, for that matter, Σ) in place of 'Herbie'. You'll find that most such substitutions yield uninterpretable gibberish—i.e., expressions that can enter into no inferential relations with the antecedently interpreted members of Σ. For instance, substituting 'jump' for 'Herbie' yields 'David admires jump', which has no inferential consequences. Same for 'jumps rapidly', 'red', 'we', and …'rential cons'…. By contrast, a commitment to 'Colorless green ideas sleep furiously' would presumably preclude entitlement to 'Nothing ever sleeps furiously', 'There are no colorless green things', 'Ideas can only be red', and many other propositions. There is a clear sense, then, in which this famous sentence is perfectly well interpretable. (It's even false!)

Setting aside gibberish, there will be a subclass of expressions that, when substituted for 'Herbie' in 'David admires Herbie', yield *interpretable* sentences, such as 'David admires Jessica', 'Jessica admires Herbie', 'David feeds Herbie', and 'David feeds Jessica'. (Again, a sentence is interpretable just in case it can play the role of premise or conclusion in an inference.) This subclass of Γ, call it Π, contains all and only the recombinable elements—i.e., the subsentential units of the language—including words, phrases, clauses, morphemes, subjects, predicates, or whatever other syntactic categories the language in question contains. We can now call one sentence, S, a *substitutional variant* of another, S*, just in case S is the result of substituting one element of Π with another in any member of Σ. Thus, 'David admires Herbie' is a substitutional variant of 'Jessica admires Herbie', on account of its being the result of the substitution of 'David' for 'Jessica'.

The foregoing puts us in a position to entertain a new inferential relation between sentences. Let's call an inference *substitutional* just in case the conclusion is a substitutional variant of one of the premises. The two inferences, from 'David admires Herbie' to either 'David feeds Herbie' or to 'David admires Jessica', are both fine examples. This no-

[12] Any *actual* creature's primary linguistic data (PLD) will, of necessity, be finite for in the course of language acquisition. The obvious analogy to the case of language acquisition in *human* children should not tempt us into assuming that Brandom is pitching an empirical account of the stages of acquisition. Still, the analogy is worth noting, even if we strongly suspect—as generativists do, *pace* Tomasello (2005)—that children's linguistic capacities are productive/generative *right from the get-go*. From the latter hypothesis, it follows that there is no such thing, really, as a finite set of PLD for the child, the child's acquisition device is *always* doing something analogous to hypothesis testing, *even in the absence of input data*. On this picture, the set of PLD is a constantly-moving target—in effect, a massively complex mental representation, or representational structure/system, within the child. The latter is plainly *not* identical with the set of utterances that happened to be produced in a child's presence.

tion of a substitutional inference is what allows for an application of the inferentialist strategy to subsentential expressions.

Subsentential Meaning: The meaning of a subsentential expression, α, is the set of *materially good substitution inferences* involving α.

Thus, the meaning of 'Herbie' is the set of inferential proprieties that includes {'David feeds Herbie' ↔ 'David feeds Herb'}, {'David feeds Herbie' ↔ 'David feeds a dog}, and {'David feeds Herbie' ↔ 'David feeds *his* dog'}, and many others.[13] In all such cases, the substitutional inferences are *materially good* in virtue of the fact that 'Herbie' is substituted by any of his other actual names, or by other ways of correctly describing him, uniquely or otherwise.

Needless to say, no one—not I, and certainly not Herbie—will ever have a *full* grasp of the set of inferential proprieties that governs the use of the expression 'Herbie', as this would involve knowing everything there is to know about him. Nor is there any guarantee that any two speakers will converge from the outset on what is correct to infer from "David feeds Herbie"—e.g., whether inferring "David feeds a dog" is (materially) good. Rather, the point is this: given that there are, in point of fact, *plenty* of ways for me to entitle myself to "Herbie is a dog", and no plausible ways (please grant) to undercut that entitlement, it would be *incorrect*, pragmatically *improper*, and epistemically *unwarranted* for someone to assert the opposite. This holds even if my interlocutor is *strongly disposed* to maintain a contrary position on the matter (foolishly, no doubt). It's important to always keep in mind that *normative* inferentialism is not about inferential *propensities*; it's about inferential *proprieties*.

1.6. *Predicates and singular terms*

One consequence of the view presented thus far is that some linguistic expressions can be inferentially stronger or weaker than others. Consider the verbs 'runs' and 'moves'. The latter is logically stronger than the former because all substitution inferences from '*x* runs' to '*x* moves' are good, but the reverse inferences generally aren't. In such cases, the substitution inferences are said to be *asymmetric*. We also find terms that *invariably* enter into *symmetric* substitution inferences—e.g., from 'Mark Twain was an American' to 'Samuel Clemens was an American' and back again. To make the latter type of inference explicit, subsentential expressions of identity and nonidentity can be introduced, yielding propositions of the form α=β and α≠β (e.g., 'Sam Clemens is identical with Mark Twain' and 'David is not Herbie').[14]

[13] For simplicity of presentation, I suppress issues to do with possessives like 'my' and 'his', and indexical expressions more generally. Brandom (1994, 2008) supplies an account of these, but the details are irrelevant here.

[14] The notion of "introduction" that I intend here is the one developed in Brandom (2008). Roughly, a community is capable of introducing a novel expression,

As we will see in §4, Brandom holds that the distinction between predicates and singular terms comes down to the distinction between those expressions that *must* license *only* symmetric inferences (e.g., 'Herbie' and 'the dog'), and those that merely *can* license symmetric inferences, but *need not* do so (e.g., 'deer-like', 'jumps', and 'rapidly). On the basis of this claim, Brandom goes on to develop a complex line of reasoning whose ultimate conclusion I'll call the "asymmetry constraint".

Asymmetry Constraint: Any language that draws no distinction between predicates and singular terms (conceived in the above manner) *is in principle precluded from introducing conditionals*—i.e., expressions that make explicit one's commitment to the goodness of an inference—and other basic operators of propositional logic.

This claim will come to the foreground when we contrast it with Pietroski's predicativism, according to which there are in fact no singular terms at all in natural languages. If Brandom's argument succeeds, then Pietroski's predicativist semantic theory faces a serious challenge. Contrapositively, if Pietroski's predicativism is correct, then there must be a flaw in Brandom's reasoning. This is, in fact, the final puzzle for the overall reconciliation project that I'll be urging here.

1.7. *Types, tokens, and anaphoric chains*

The expressions discussed thus far have all been linguistic *types*, tokens of which may well diverge in meaning from their primary significance in the language. Indeed, terms like 'Herbie' have so many different uses—one for my dog, another for the pianist, Herbie Hancock, and countless others—that Brandom needs an account of what makes any use of 'Herbie' *semantically co-typical* with any other. The question applies even to *intra*-sentential occurrences: What makes it the case, for instance, that both tokens of 'Herbie' in 'Herbie admires Herbie' of the same type in a given communicative context?

In providing his answer, Brandom introduces the last of the major technical notions that he needs in order to carry off his overall project—viz., the notion of *anaphora*. Linguists and philosophers have paid a great deal of attention to *intra*-sentential anaphora, as in 'If a man is a

in this sense, just in case its members already have the *practical* abilities that are necessary and sufficient for being able to express—i.e., to make explicit—normative attitudes that were previously only implicit in their practice. Thus, the practical ability to implicitly treat someone as having entitled themselves to q by committing themselves to p is both necessary and sufficient to introduce conditional expressions that make explicit the material goodness of that inference—e.g., '$p \supset q$' and 'If p then q'. We will see in §2.5 that Pietroski's notion of concept introduction is different from Brandom's, and arguably orthogonal.

police officer, then he was born out of wedlock', where the pronoun 'he' is anaphoric on 'a man'. Syntacticians, in particular, have devised principles of generative grammar that aim to explain the natural distribution of anaphoric expressions within sentences of natural language. Somewhat less effort has thus far been expended on analyzing *inter*-sentential anaphora, as in the following exchange between speakers Mihir and Rushal.

> Mihir: That man seems to have fallen ill right after he approached the police line.
> Rushal: He must have gotten hit by their fancy new sonic weapon.
> Mihir: Oh, hey, I didn't see you there! Do you happen to know the guy?
> Rushal: No, I just heard you talking about him and I figured I'd chime in.

Here, an anaphoric chain is initiated by Mihir's use of 'That man', which is then picked up by 'he' later in the same sentence. But the chain doesn't end there. Rushal's use of 'He' is anaphoric on Mihir's use of 'That man' and 'he'. Mihir's response picks up the anaphoric chain with an occurrence of 'the guy', which then continues onward to Rushal's use of 'him', and to occurrences of other expressions in subsequent discourse. Setting aside syntactic issues, what can we say about this phenomenon at the level of *meaning*?

In keeping with his inferentialist semantics, Brandom argues that an anaphoric chain is one in which the inferential proprieties governing the anaphoric initiator (e.g., Mihir's use of 'That man') are *inherited* by subsequent expressions in the chain. Thus, if Mihir's use of 'That man' is partly governed by his commitment to 'That man is falling ill on live television', then Rushal inherits this commitment (among others) in picking up the anaphoric chain with the use of 'He', along with whatever entitlements for this claim Mihir had already secured prior to Rushal's appearance on the scene.

With this account in hand, Brandom treats as a special case occurrences that are treated as semantically co-typical *because* they are phonologically or orthographically co-typical—e.g., the two occurrences of 'Herbie' in 'Herbie admires Herbie'. From this perspective, *all* expression types consist of long-stretching anaphoric chains of individual use—an idea familiar from causal theories of reference-borrowing, though shorn of various optional commitments. This account also makes it clear what's happening at the level of *pragmatics*. In picking up anaphoric chains, speakers are able to take on normative statuses—paradigmatically, commitments and entitlements—without themselves having *explicitly avowed* those statuses, and often without having much (if any) idea what exactly it is that they've inherited. To illustrate, we can extend the above example.

Suppose Rushal had no prior commitments regarding the victim's appearance on television, or indeed anything at all about the victim, but was strongly committed to the claim that police don't use sonic weapons on camera. In that case, upon being subsequently apprised of Mihir's entitlement to 'That man is falling ill on live television', Rushal will be under normative pressure to either revise his prior commitments about on-camera police violence, or to withdraw the claim that the victim *must* have been affected by a specifically *sonic* weapon. In this second case, the revision can target either the predicate 'sonic'—perhaps the police used an invisible gas—or the alethic modal expression ,'must'. The latter, on Brandom's view, functions to make explicit the *modal robustness* of an inference—i.e., its insensitivity to substitutions of background auxiliary commitments, up to some boundary conditions (e.g., physical law). In the present case, the boundary conditions are set by Rushal's commitments regarding the general institutional practices of local police. In order to regain epistemic equilibrium, Rushal can revise various commitments concerning these practices; for instance, he might conclude that the local Sheriff has deemed this to be a special occasion, on which on-camera use of sonic weapons is warranted.

1.8. *Summary*

We've now surveyed the main contours of Brandom's overall philosophical project. The explanatory strategy he pursues can be characterized as "top-down", in the sense that he begins by offering an account of communal normative practices, in the broadest sense, and identifies within these an important subclass—namely, practices that serve to institute distinctively *linguistic* norms governing assertion and other communicative acts. (One last plea for demands!) Such norms pertain to the inferential proprieties that expression types have in their semantically primary occurrences. Thus, the account moves "down" a step—from a normative pragmatics that posits statuses of commitment and entitlement, to an inferentialist semantics that aims to analyze *meaning* in *terms of* these statuses. The meaning of a propositional expression type is, on this picture, identified with its normative inferential role—i.e., what *other* claims it commits or entitles one to, and what commitments one must undertake in order secure an entitlement to it.

Drilling down still further, Brandom develops the substitutional approach, which allows one to "dissect" proposition-sized expression types, revealing subsentential bits of vocabulary. These carry their own "ingredient content", despite lacking the *free-standing* significance of propositional expressions that enter directly into inferences as premises or conclusions. The details of this proposal put in place the theoretical commitments that Brandom needs in order to distinguish predicates from singular terms—a distinction that he goes on to argue will be discernable in *any* linguistic practice that allows for the introduction of conditionals and other logical operators (§4.3).

Having offered a treatment of propositional and subsentential expression types, Brandom steps down another rung on the explanatory ladder, developing a conception of *anaphora* that applies far more broadly than standard discussions in the literature might lead one to suspect. The anaphoric relationship is, on this view, one of *inferential inheritance*, wherein the proprieties governing the use of one expression—the initiator of an anaphoric chain—are taken to then *also* govern the expressions occurring later in the chain, irrespective of the speaker's acknowledgement (or even awareness) of the statuses they've thereby undertaken. The latter condition serves to explain how speakers can felicitously use expressions whose total set of inferential proprieties is *unknown* to them, and perhaps even to *anyone* in the community.

One might think that all of this is utterly wrongheaded right from the get-go—the normativity, the substitutions, and even the top-level goal of delineating language-use *as such*. Indeed, from the perspective of a mainstream contemporary linguist or philosopher of language, Brandom's whole "top-down" explanatory strategy will seem downright perverse. The more common *bottom-up* alternative goes as follows.

Taking for granted the notions of denotation/reference and satisfaction, as applied to subsentential expressions, the bottom-up theorist seeks to formalize a compositional apparatus for building propositions out of them. Free-standing propositional complexes are thereby recursively assigned their own special kind of semantic value: e.g., possible-worlds truth conditions (Heim and Kratzer 1998) or sets of possible worlds (Stalnaker 1984). This, in turn, opens the door to a theory of linguistic *communication*, according to which speakers append illocutionary forces to the range of recursively-specified meanings, yielding a variety of speech-act types (questions, commands, etc.). The inferences in which a (now-interpreted) speech act type figures can then be classified as good or bad in virtue of the semantic structures that the combinatorial apparatus assigns to their premises and conclusions, as well as the illocutionary forces that (somehow) "attach" to those structures.

Having thus analyzed the semantic properties of speech acts and inferences, one might note that some—perhaps, in the end, *all*—of these have features that reliably trigger unencapsulated pragmatic reasoning. This motivates the familiar project of supplementing a pragmatic theory with "maxims" of rational cooperative communication/action (Grice 1989; Sperber and Wilson, 1986). Theorists who have carried out this latter project have developed impressive accounts of implicature, metaphor, and other complex communicative phenomena (Levinson 1983; Harris 2020).

Proponents of the bottom-up strategy have pressed a catalogue of objections to Brandom's project. These include, but are not limited to, the following: (i) insistence on a *compositionality* constraint that the inferentialist allegedly can't accommodate; (ii) rejection of the idea that language is fundamentally a *communicative* system; (iii) requirement

that any *legitimate* inquiry foreswear trafficking in normative assessments; and (iv) an allegation to the effect that normative inferentialism is incompatible with what is known empirically about the human mind/brain, particularly in respect of its language-processing abilities.

Before any of these challenges can be met, each stands in need of careful articulation. As previously noted, I believe that such a task is best undertaken by pitting Brandom's project against what appears, at first blush, to be a rival alternative. (As advertised, I'll argue afterwards that the appearances are often deceiving in this regard.) With that in mind, I now turn to the work of Paul Pietroski, whose semantic theory is a recent and powerful contribution to the larger enterprise of generative linguistics.

2. *Pietroski: Meanings as pronounceable instructions for concept assembly*

The theoretical commitments that comprise Paul Pietroski's approach to natural-language semantics are advanced and defended in his recent book, *Conjoining Meanings* (henceforth *CM*).[15] In this section, I summarize several of Pietroski's main contributions, highlighting aspects of his view that bear on my ecumenical strategy in §§3-4. To be clear from the outset, the ideas laid out in *CM* strike me as constituting genuine progress in our understanding of the psychological mechanisms of human language use. Moreover, I find wholly compelling his arguments against the central pillars of received views in semantics—particularly, the commitment to an extensional/truth-conditional approach. The book, overall, is replete with rich and instructive discussions of topics that go well beyond the scope of the present discussion. But while we won't be able to look at the details of some of Pietroski's original proposals here, it's worth noting that they are all, to my mind, persuasively motivated by historical, formal, and empirical considerations. That having been said, let's dive in.

2.1. *A different methodology and new explanatory aims*

While Brandom's inferentialist approach is virtually unknown in cognitive science, the methodology of generative linguistics will be familiar to many in the field, at least in broad outline. Rooted in a foundational commitment to *naturalistic* inquiry, the idea is to treat language as a biological phenomenon—not necessarily in the sense that it has an adaptive function (Chomsky [2016] disputes this), but in the sense that a neurophysiologically realized cognitive structure is the explicit target of inquiry. The linguist thus works on the assumption that human minds contain a language-specific device—a "faculty", "module", or "mental organ"—with a distinctive computational architecture, a

[15] This section elaborates the material in Pereplyotchik (2019). The operative notion of a *subpersonal* level of description is spelled out in Pereplyotchik (2017: ch. 7).

proprietary representational format, and dedicated/domain-specific information-processing routines. The goal is to provide a detailed specification of each of these, yielding a neurocognitive account of the acquisition and use of language.

On analogy with bodily organs, the faculty of language (henceforth FL) is assumed to "grow" within the child during the early years of development. This happens in accordance with a genetic program, phenotypically realized in the child's innate ability to acquire linguistic competence under a diverse range of social and environmental circumstances. Thus, a central aim of generative linguistics is to specify not only the grammar of an adult language, but also the principles that underlie language *acquisition*—particularly those that allow the child to home in on a *specific* grammar in a relatively short time, with little or no (overt) negative evidence (Chomsky 1986; Yang 2006). This problem is made exceedingly challenging by the fact that natural languages are invariably productive/generative, meaning that they allow for boundless applications of combinatorial recursive operations, yielding a discrete infinity of nonredundant[16] interpretable structures.

The generativist's strategy for dealing with this central feature of natural language is to posit grammatical principles that are inherently *compositional* at all levels of analysis—phonology, morphology, syntax, and semantics. The syntactic module of FL is taken to merge the elements of the lexicon—atomic units of a language that contribute their distinctive meanings to more complex structures. On the basis of these, the semantic module recursively generates complex *meanings*, which can enter into downstream personal-level cognition—judgment, reasoning, planning, and the like.[17]

Pietroski's main goal in CM is to characterize the semantic module by offering a detailed proposal about its proprietary representational format—specifically, the nature of the lexical items—and the computational operations that assemble larger interpretable structures. At the level of format, the hypothesis he develops is that virtually all lexical items are predicates, the latter being restricted to only two types—monadic and (semi)dyadic. Regarding computational operations, Pietroski aims to make do with a bare minimum of compositional semantic principles, with the lion's share of work being done by nothing more than two flavors of predicate conjunction (one for each type of predicate).

[16] Pietroski points out that this goes well beyond mere recursion, which is trivially satisfied by any languages with a rule for applying sentential operators. The infinitude of English thus differs *qualitatively* from the infinitude of a language that permits the formation only of P, P&P, P&(P&P),… or P, ~P, ~~P, ~~~P,….

[17] It's important to note that what has been said thus far is not (yet) intended as a theory of real-time/on-line language processing. Rather, it is to be seen as an abstract characterization of the architecture and internal operations of a specific cognitive structure, acquired at birth and persisting in a stable state thereafter (Chomsky 1995).

We'll look at some of the details shortly, maintaining our present focus on matters of methodology.

Following Chomsky (1986, 1995, 2000), Pietroski adopts an *individualist* position, taking the object of study to be an "I-language"—an *i*ntensionally-specified procedure *i*nternal to an *i*ndividual language user. He supports this with forceful arguments against the alternative conception of language(s) that we find in the work of David Lewis (1969, 1970, 1973). A language, on this rival picture, is a kind of abstract object—namely, an extensionally-specified a set of well-formed sentences—which is "selected" by a population of creatures, via the adoption of social/communicative "conventions". The latter Lewis sees as jointly constituting a *public* language, such as English or Norwegian—what generativists refer to as "E-languages". Pietroski rejects virtually every aspect of this picture. We'll look at his reasons for doing so in §4. For now, it's sufficient to distinguish three key points of contention.

First, there's the metaphysics. Lewis (1973) says languages are abstracta, whereas Pietroski sees them as biologically-instantiated computational procedures. Then there's the issue of extensionality. Pietroski rejects Lewis's theoretical goals, which consist merely of extensionally specifying meaning-pronunciation pairs, and adopts instead a more weighty explanatory aim—namely, that of specifying human linguistic competence as a function-in-*intension*. Only in this way, he argues, can the resulting theory capture the *psychologically real* operations that yield interpretable structures. Finally, there's the issue of publicity, and related troubles with Lewis's notion of "selection". Pietroski's individualist stance leads him to eschew the folk-ontological commitment to public languages, at least for the purposes of mature empirical inquiry. This manifests in his methodological practice of focusing on matters of individual psychology—e.g., internal mechanisms of semantic composition—rather than the social practices of linguistic communication. Accordingly, Pietroski sees Lewis's appeal to public conventions as generally unhelpful for—indeed, an outright *distraction from*—the empirical study of linguistic meaning.

Pietroski's disagreements with Lewis go well beyond such methodological issues, extending to matters of technical detail. For, in addition to the large-scale commitments mentioned thus far, Lewis (1970) also developed a powerful formal apparatus for conducting semantic theorizing. Expressions, in this scheme, are assigned "semantic types", which are either basic or recursively derived. The interpretation of complex structures is then accomplished by functions that map one semantic type onto another. In its most familiar version, such a semantic theory will assign sentences the basic type <t> and singular terms the basic type <e>. Thereafter, monadic predicates can be treated as having the derived type <<e>, <t>>, which is a function from things of type <e> to things of type <t>.

Although this formal typology presupposes no particular metaphysics or metasemantics, it's common in practice to think of singular terms

as denoting entities (e.g., Jessica), and sentences as denoting truth-values (T and F). With this in place, monadic predicates like 'swims' can be assigned the semantic function of mapping the entities in its domain to the truth-values in its range. For instance, 'Jessica swims' is mapped to T just in case Jessica (the actual person) satisfies the predicate 'swims'; otherwise, F. Likewise, adverbs such as 'often' and 'expertly' have the derived type <<<e>, <t>>, <t>>, which is a function that maps the semantic value of predicates (i.e., functions from <e> to <t>) to the semantic values of sentences (i.e., T or F). Put somewhat imprecisely, the intuitive idea is that 'Jessica swims expertly' is mapped to T just in case 'expertly' is satisfied by the predicate 'swims' when applied to 'Jessica'.

It's no exaggeration to say that this general framework is seen as a foundational contribution to formal semantics, even by generative linguists who have no truck with—or, indeed, no awareness of—Lewis's broader projects. Part of what makes Pietroski's negative contentions so radical, then, is that he rejects wholesale this now-mainstream approach to semantic theorizing. In particular, he argues that taking an infinite hierarchy of types as explanatorily primitive is not only unparsimonious, but leaves wholly unexplained crucial aspects of the natural languages that children invariably acquire. As a matter of empirical fact, humans language permits the construction of only a *limited* class of semantic types, not the infinite range of logically possible ones. This empirical generalization plainly stands in need of explanation, which a semantic theory can't provide if it takes all possible types as available to a speaker right from the start.

> One can say that thinkers must have the requisite abstractive powers, given the capacities required to form thoughts like ABOVE(FIDO, VENUS) & BETWEEN (SADIE, BESSIE, VENUS). But one needs an account of these alleged powers—which permit abstraction of a tetradic concept from ABOVE(_, _) and BETWEEN(_, _, _)—to explain how thinkers can form the concepts that *Begriffsschift* expressions reflect. This is not to doubt the utility of Frege's logical syntax. On the contrary, his proposals about the architecture of thoughts were major contributions. But Frege insightfully invented a logical syntax whose intended interpretation raised important questions that he did not answer.
>
> One can insist that given any polyadic concept with n unsaturated "slots," a human thinker can use n-1 saturaters to create a monadic concept, leaving any one of the slots unfilled. But that leaves the question of how we came to have this impressive capacity. And in chapter six, I offer evidence that a simple form of conjunction lies at the core of unbounded cognitive productivity. Our natural capacities to combine concepts are impressive, but constrained in ways that suggest less than an ideal Fregean mind.

Pietroski recommends a more parsimonious alternative—one that eschews the infinite hierarchy of semantic types and posits only a very small handful, including, most importantly, monadic and quasi-dyadic predicates. "The idea [is] that with help from Slang syntax, we can generate an analog of GIVE(VENUS, _, BESSIE) without saturating GIVE(_, _, _)

—much less saturating it twice, or thrice, and then desaturating once" (103).

Nor does his iconoclasm end there. As noted earlier, Lewis's general framework for semantic theorizing leaves open a variety of issues in metaphysics and metasemantics. An equally mainstream approach to natural-language semantics is decidedly more committal on these points. Donald Davidson's truth-theoretic semantics (Davidson, 1983), as well as the many variants of it that have now been developed, identifies the meanings of linguistic expressions with their *extensions*. Thus, truth conditions (perhaps relativized to possible worlds) are seen as the semantic values of sentences; entities are the values of singular terms; sets are the values of predicates; events in the case of verbs, and so on. Pietroski marshals a battery of arguments against this familiar approach. We'll examine these shortly. For now, we note only that this anti-extensionalism is a core commitment that he shares with Brandom. It is, therefore, a major plank in the bridge that I aim to build between the two in §§3-4.

2.2. *Meanings are definitely not extensions*

Pietroski sees semantics as a naturalistic inquiry into "how Slang expressions are related to human concepts" (115). Some theorists wish to simply *identify* meanings with concepts, but Pietroski points out that this leaves wholly unexplained the psychological processes that *constitute* our semantic competence. I'll argue in §4 that this point applies to Brandom, who sometimes speaks indiscriminately of meanings, concepts, conceptual contents, intentional contents, discursive contents, propositional contents, and so on. However, as I'll emphasize there, the difference can only be viewed as a substantive theoretical *dispute* if we let their use of the folk term 'meaning' bewitch us into assuming that they have a common explanatory target, contrary to fact.

Better, I think, to appreciate the highly theoretical nature of this piece of jargon and the different—*but not thereby incompatible*—explanatory goals of the two frameworks in which it shows up. Thus, we can distinguish meaning_B from meaning_P and proceed to contemplate how the two are related, this now being a jointly philosophical and empirical question, not a boring verbal one. Indeed, this point is made explicitly by both Pietroski and Brandom, in connection with both 'meaning' and another vexed notion—that of 'concepts'—which notoriously plays a wide variety of roles in diverse research contexts. Here again, we can speak of concepts_B and concepts_P, aiming to articulate the relations between them. Likewise for 'thought', 'judgment', and other terms, when explicit disambiguation is required. (See also the discussion of the notorious '-ing/-ed' ambiguity in §4.2.)

As noted above, another popular idea is to identify meanings with extensions (Davidson 1983). The central negative contention of *CM* is that the notions of extension, truth, and denotation should play *no ex-*

planatory role in a psychologically-oriented semantics for natural languages ("Slangs"). Pietroski argues persuasively that the best empirical theory of the relation between Slang expressions and concepts will *not* identify meanings with extensions. Indeed, he rejects even the weaker claim that meanings *determine* extensions. He proposes, instead, to identify meanings with something entirely different—in particular, something that can play the psychological role of relating language to cognition. The candidate he recommends is this: *pronounceable instructions for accessing and assembling concepts*. We'll look at this in some detail, but let's first get clear on *why* Pietroski rejects the truth-conditional orthodoxy that dominates formal semantics. As we'll see, there are a great many reasons. To my mind, no one of these is necessarily decisive, but, taken together, they strongly suggest turning away from the extensionalist project and starting anew, *however* much revision this might require. As we go along, I'll land a few jabs of my own.

2.3. *Objections to truth-conditional semantics*

Pietroski views truth-conditional semantics (henceforth 'TCS') as an empirical hypothesis about Slang expressions, according to which there is a relation—call it "true of", "refers to", "denotes", or whatever you like—that holds between words and items in the world. TCS views this relation as being of central importance to our theoretical characterization of natural-language meanings. In rejecting this hypothesis, one need not deny, of course, that there are words or that there objects (e.g., babies and 'bathwater'). One can, instead, deny that there is a unique relation between them, let alone one that's suited to playing the theoretical role of *linguistic meaning*. Here is how Pietroski puts the point:

> I don't think 'sky' is true of skies (or of sky), much less blue skies or night skies. I don't deny that there are chases, and that in this sense, chases exist even if skies don't. But the existence of chases doesn't show that 'chase' is true of them… [Likewise], there is no entity that 'Venice' denotes. In this respect, 'Venice' is like 'Vulcan', even though one can visit Venice but not Vulcan… I also agree that there is a sense in which there are blue skies, but no blue unicorns. But it doesn't follow that 'sky' is true of some things, at least not in the sense of 'true' that matters for a theory of truth… [T]here is no call to quantify over skies, in physics or linguistics. (68)

As the example of 'Vulcan' illustrates, words can perfectly well be meaningful without having extensions. Pietroski's view is that this holds of *all* Slang expressions. What's interesting about words like 'Vulcan' is they "illustrate the general point that words don't *have* extensions". The idea isn't merely such terms have *empty* extensions; it's that they have *none at all*.

Even if words *did* have extensions, the latter couldn't be identified with meanings, if only because "expressions with different meanings can have the same 'extension'" (15). Fans of TCS will typically appeal to "non-actual possibilities" in dealing with this issue. For instance, 'unicorn' and 'ghost' are said to have the same extension in the *actual*

world, but they differ in meaning—the reply goes—because they have *different* extensions in *other* possible worlds. Pietroski correctly points out that this "is an odd way to maintain that meanings are extensions."

> If the meaning of a word is not whatever set of things that the word happens to be true of, why think the meaning is a mapping from each possible world *w* to whatever set of things that the word happens to be true of at *w*? [If] Slang expressions need not connect pronunciations to actual things, it seems contrived to insist that these expressions connect pronunciations to possible things… [I]nvoking possible unicorns is contrivance on stilts. (12)

Doubtless, fans of TCS will see this as little more than an *ad hominem*. We'll look at stronger arguments shortly. For now, I want to emphasize that this point—or, in any case, a version of it—carries more weight than is commonly appreciated. Let me take a brief aside to develop it in my own terms.

The intuitive considerations that motivate TCS (e.g., for introductory semantics students) almost always have to do with objects that are available for perceptual inspection. ('David' refers to *this guy*, 'my desk' refers to *that thing*, and so on.) This serves to illustrate, at the level of pre-theoretical intuition, how linguistic expressions "hook onto the world"—namely by way of perceptual contact (indeed, literal contact, in the case of haptic perception). Shortly thereafter, the details of one or another formal theory are introduced, giving the student little time to reflect on how far the initial illustration can plausibly generalize. (Spoiler alert: not very far!) If philosophical questions happen to arise about the status of these "reference" and "correspondence" "relations"—e.g., with regard to empty names and predicates ('Vulcan', 'unicorn', etc.)—the instructor can use the opportunity to explore various technical proposals for dealing with such "special cases"—e.g., Russell's theory of names as disguised descriptions, or the formal apparatus of possible-worlds semantics. Attention is thus deflected away from how massive the intuitive problem really is. Here's a much-needed corrective.

Consider for a moment the vast range of expressions that we can readily produce and comprehend, and reflect on how vanishingly few of these have anything much to do with what's going on in physical reality, let alone with things that we can perceptually inspect in any intuitive sense. We speak of Santa and his elves, gods and demons, goals and fears, opportunities and temptations, aliens and chem-trails, reptiles and unicorns, futures and fictions, numbers and functions, nouns and verbs, fonts and meanings, haircuts and field-goals, stocks and derivatives, mergers and monopolies, economies and governments, boson fields and spin-foams, black holes and electrons, Blacks and whites, Jews and Frenchmen, London and Moscow, classes and genders, protests and stereotypes, jocks and nerds, bits and bytes, poems and operas, humor and beauty, and even the possibility (albeit dim) of true liberatory justice.

Appreciating the sheer scope of the phenomenon to be explained renders, to my mind, *utterly implausible* the strategy of taking direct perceptual contact with the world as our model of how language relates to reality *in general*. Moreover, the total lack of convergence that we find amongst metasemanticists when we go looking for a metaphysical account of truth and reference—conceived of, again, as a Very Special sort of natural relation—strikes me as further grounds for abandoning the project of extensional semantics immediately and forthwith. It helps, of course, that Pietroski supplies a powerful *alternative* framework for doing semantics. And it certainly doesn't hurt, that Brandom complements this with an independently attractive ("deflationist") account of truth and reference.

All that aside, Pietroski has a further, more powerful argument against invoking non-actual possibilities for the purpose of individuating meanings. He makes use of Kripke's contention that the non-existence of unicorns in the actual world implies their non-existence in *all other* possible worlds (Kripke, 1980). Of course, there may well be creatures in other possible worlds that *look* a lot like what we imagine unicorns would look like. But they would not thereby *be* unicorns, and our word 'unicorn' would not thereby be true of them. If that's correct, then 'ghost' and 'unicorn' aren't just co-extensive in *our* world; they're co-extensive in *every* possible world. Thus, no identification of meanings with extensions, actual or possible, will distinguish the meanings of those two words. Likewise for all of the related cases—empty names, defective predicates, necessary falsehoods, and so on.

One might reply by rejecting Kripke's semantic and metaphysical assumptions, and adopting instead a Lewisian counterpart theory, but Pietroski points out several problems for this strategy as well. Adopting the terms 'LUNICORN' for Lewisian unicorn-lookalikes and 'KUNICORN' for the whatever it is that Kripke has in mind, he makes the following powerful retort.

> We can grant that some theorists sometimes use 'unicorn' to express the technical concept LUNICORN. But if 'unicorn' can also be used to express the concept KUNICORN, then it seems like contrivance to insist that the Slang expression has a meaning that maps some contexts onto the extension of LUNICORN and other contexts onto the extension of KUNICORN. If we assume that words like 'possibly' have extensions, then perhaps we should specify the meanings of such words in terms of a suitably generic notion of world that allows for special cases corresponding to metaphysical and epistemic modalities; cp. Kratzer. But in my view, theorists should not posit (things that include) unicorns in order to accommodate correct uses of 'Possibly/Perhaps/Maybe unicorns exist' or 'There may be unicorns'; and likewise for squarable circles.

Thereafter, the dialectic turns to matters that we need not enter into here. Suffice it to say that, even if this worry about fine-grained meanings can ultimately be defused, TCS would still face Pietroski's more technical (and potentially more damaging) objections. These include

matters pertaining to liar-sentences, as well as the more widespread and natural phenomenon of event descriptions. These too go beyond the scope of our discussion. One argument that I do want to say a bit more about, though, is on the topic of polysemy, where Pietroski's view of the matter finds wide acceptance among generative linguists—though, notably, not philosophers of language (see, e.g., Michael Devitt's paper in this issue.)

Following Chomsky (2000), Pietroski points out that 'water' is polysemously used to talk about many substances—those found in wells, rivers, taps, etc.—nearly all of which have *lower* H_2O contents than substances that, at least *prima facie*, are *not* water, including coffee, tea, and cola (*CM*, 21). This presents a challenge to theories that view 'water' as bearing a reference relation to (all instances of?) the natural kind *water*, whose metaphysically essential property is *being composed of H_2O molecules* (Kripke, 1980). If coffee, tea, and cola all have more H_2O in them than most ordinary instances of water, then it's not clear why 'water' doesn't bear the reference relation to *them*, rather to the stuff in the local rivers and wells.

A related consideration has to do with predicate conjunction. The word 'France' can be used in expressing either of two concepts: FRANCE:BORDER and FRANCE:POLIS. The border is hexagonal and the polis is a republic. But, Pietroski points out, the polysemy of 'France' "does not imply that something is both hexagonal *and* a republic, much less that 'France' denotes such a thing" (74). Similarly, while 'London' can be used to talk about "a particular location or a polis that could be relocated elsewhere," it is plain that "no location can be moved, and no political institution is a location." Pietroski concludes that "no entity is the denotation (or 'semantic value') of 'London'; the ordinary word *has* no denotation" (73, emphasis mine).

2.4. *Meanings as pronounceable instructions*

Let's turn now to Pietroski's positive views. As noted earlier, the main goal of *CM* is to defend the hypothesis that linguistic meanings are "pronounceable instructions for how to access and assemble concepts" (1). More specifically,

> each lexical meaning is an instruction for how to access a concept from a certain address, which may be shared by a family of concepts. … A Slang expression Σ can be used to access/build/express a concept C that is less flexible than Σ—in terms of what Σ can be used to talk about, and how it can combine with other expressions, compared with what C can be used to think about and how it can combine with other concepts— since Σ might be used to access/build/express a related but distinct concept C* .

Unpacking Pietroski's hypothesis requires getting clear on the three key notions of *pronounceable instructions*, *compositional assembly*, and *conceptual types*. Each is more challenging than the last, so we'll start with instructions and work our way up.

2.4.1. *Pronounceable instructions*

An utterance of a sentence is a spatiotemporally located event, in which a speaker produces a physical signal. The latter serves, on Pietroski's view, as an instruction for the hearer's FL to perform a computational procedure.[18] The instruction can be carried out by any hearer whose I-language is sufficiently similar to the speaker's. The acoustic properties of an utterance, upon being transduced, trigger an early perceptual constancy effect, whereby a dedicated module imposes phonological categories on the neural encoding of the acoustic blast. These cognitive operations serve, in turn, as instructions for the further segmentation of the phonological units into syllables and eventually into morphemes and other lexical items. The latter, on Pietroski's view are best seen as instructions for accessing ("fetching") individual concepts, which he conceives of as atomic units of one or another language of thought. I say "one or another" because his view leaves open the possibility, which he goes on to explore and even endorse, that there are *many* languages in which the mind conducts its information-processing. We'll return to this point in connection with Pietroski's discussions of Frege (§2.5).

Importantly, Pietroski maintains that concepts reside in semantic "families", which have their own "addresses" in a broader cognitive architecture. This is a large part of his explanation of the aforementioned phenomenon of polysemy. The idea is that one and the same lexical item can be an instruction for fetching "a concept from a certain lexical address … shared by a *family* of concepts" (8). Because a lexical instruction points only at an address, rather than a specific concept, it's left open for *downstream* processing routines to determine which particular concept from the indicated address/family is "relevant" in the present context.

This, of course, raises deep and difficult questions about how hearers manage this latter step—i.e., reliably accessing the relevant concept(s) in a given context, rather than the irrelevant ones from the same conceptual family. What psychological mechanisms select just *one* of a family of concepts residing at a common lexical address? In large part, Pietroski leaves this issue open—justifiably so, given everything else he's juggling. But it's worth remarking in the present context that the mechanisms of this kind of selection are widely agreed to involve—indeed, to *require*—precisely the kind of nondemonstrative pragmatic reasoning that Brandom has argued to be constitutive of conceptual contents.

[18] "I hope the analogy to elementary computer programs, which can be compiled and implemented, makes the operative notion of instruction tolerably clear and unobjectionable in the present context. … Instructions can take many forms, including strings of '0's and '1's that get used—as in a von Neumann machine—to access other such strings and perform certain operations on them. … And instead of arithmetic operations that are performed on numbers, one can imagine combinatorial operations that are performed on concepts" (108).

2.4.2. *Assembling concepts*

Turn now to the second key notion in Pietroski's main hypothesis—viz., the compositional assembly of concepts. In general, instructions for assembling something can vary along any number of dimensions. Some are clear; some aren't. Some are detailed; others are vague. Some are simple; others are complex—i.e., composed of simpler instructions. Moreover, not everything to which an instruction is presented is capable of carrying it out. Some computers can't run the software that others can. Some chefs can't bake the cakes that others have no trouble baking. And some proteins (or cells) can follow genetic instructions that others simply can't. Lastly, the *products* of successfully carrying out instructions can vary widely. The same student, with the same instructions, can succeed or fail on an exam, depending on whether they've had sleep the night before. Likewise, a novice barista will generally make worse coffee with low-quality ingredients than with high-quality ones, successfully following the same instructions both times.

Given that the semantic module of FL is assumed to have a stable processing routines, carried out in a proprietary representational format, it follows that it won't be able to process just *any* old instruction, but only a restricted kind. Likewise, it will only be capable of assembling only a limited class of outputs. The question, then, is what kinds of instructions the semantic module is capable of implementing and what sort of structures it's capable of building.

Many theorists aim at capturing something called "compositionality"—a piece of theoretical jargon that, perhaps more than most, has been worn smooth by a thousand tongues (to use Wilfrid Sellars's clever phrase). Of the many ways of cashing it out, Pietroski maintains that what's required for an avowedly *cognitivist* project is that the meanings of lexical items compose in ways that suitably mirror the structure of complex concepts. Thus, having identified the meanings of lexical items with instructions to fetch individual concepts, he argues that these instructions compose, forming *complex* instructions, with some functioning as (detachable) components of others. These semantic instructions—what Pietroski calls *Begriffsplans*—are responsible for the assembly of concepts meet two constraints. First, they must be suited to *that specific type* of instruction. While other kinds of human concepts might be assembled by non-linguistic means, *Begriffsplans* can only assemble concepts of a very specific nature (to be spelled out shortly). Second, in keeping with the "mirroring" constraint (my word, not his), the complex concepts that *Begriffsplans* assemble must bear the same part-whole relationships to one another as do the *Begriffsplans* themselves.

Laying out some of the specifics of the *Begriffsplans* that Pietroski posits will put us in a position to better appreciate his views on concepts. The clearest case of this pertains to instructions for *predicate conjunction*. Pietroski takes this to be an absolutely central aspect of

linguistic concept assembly, in part because he holds that the kinds of concepts that the human FL is capable of assembling are uniformly *predicative*. In saying this, he means to deny outright that natural languages ("Slangs") allow us to access singular concepts. Such concepts do exist, he thinks, but they *can't* be fetched by *Begriffsplans*. Indeed, he holds that the *only* predicative concepts FL can fetch, and hence assemble, are limited to just the monadic and the quasi-dyadic, with higher adicities receiving a different analysis. These two types of concept correspond to two flavors of predicate conjunction: M-junction and D-junction. Here's how Pietroski characterizes the overall process.

> If biology somehow implements M-junction and D-junction, one can envision a mind with further capacities to (i) use lexical items as devices for accessing simple concepts that can be inputs to these operations, and (ii) combine lexical items in ways that invoke these operations. … Suppose that combining two Slang expressions, atomic or complex, is an instruction to send a pair of corresponding concepts to a "joiner" whose outputs can be inputs to further operations of joining. Imagine a mind—call it Joyce—that has some lexical items, each with a long-term address that may be shared by two or more polysemously related concepts. Joyce also has a workspace in which (copies of) two concepts can be either M-joined or D-joined to form a single monadic concept, thereby making room for another concept in the workspace, up to some limit. Joyce can produce and execute instructions like fetch@'cow'; where for each lexical item L, the instruction fetch@L is executed by copying a concept that resides at the long-term address of L into the workspace. Joyce can also produce and execute instructions of the forms M-join[I, I0] and D-join[I, I0]; where I and I0 are also generable instructions. An instance of M-join[I, I0] is executed by M-joining results of executing I and I0, leaving the result in the workspace, and likewise for an instance D-join[I, I0].

Having introduced two basic types of composable *Begriffsplans*—one for fetching concepts like DOG(_) and one for assembling these into complex structures—Pietroski adds four other types of basic semantic operation:

(i) a limited operation of *existential closure*
(ii) a mental analog of *relative clause formation* (weaker than λ-abstraction)
(iii) the *introduction* of concepts like GIVE(_) on the basis of GIVE(*x, y, z*)
(iv) of *thematic* concepts—e.g., AGENT(_), PATIENT(_), RECIPIENT(_)

> [G]iven two monadic concepts, the operation of M-junction yields a third such concept that applies to an entity e if and only if each of the two constituent concepts applies to e. (32) … In short, Slangs let us access and assemble monadic [and some limited dyadic] concepts that can be conjoined, indexed, polarized, and used as bases for a limited kind of abstraction.

We'll look at several of these operations in more detail below, but the following passage contains an initial illustration of the kinds of structures that this system can assemble.

> My claim is not that 'gave a dog a bone' is an instruction to build [just *any*] concept with which one can think about things that gave a dog a bone. That instruction might be executed by building the concept ∃y∃z[GAVE(x,

> y, z) & BONE(y) & DOG(z)], which has a triadic constituent. My claim is that 'gave a dog a bone' is an instruction for how to build an M-junction like [[GIVE(_)^PAST(_)]^∃[PATIENT(_, _), BONE(_)]]^∃[RECIPIENT (_, _) ^DOG(_)], which has only an occasional dyadic concept that has been "sealed in."

This passage usefully contrasts the conceptual structures assembled by FL with the those that are often assumed by linguists—wrongly, by Pietroski's lights—to be available to humans *antecedent* to the development of language.

2.4.3. *Concepts, predicative and sentential*

We are now in a position to ask more specific questions about Pietroski's third key notion—viz., that of a concept. As we've already seen, he takes these to be expressions in a compositional language of thought, some of which can be assembled by the semantic module of FL. But, however they might be assembled, they are the representations that allow us to think about the world.

> [C]oncepts have contents that can be described as ways of thinking about things; cf. Evans. A concept that can be used to think about something as a rabbit, whatever that amounts to, has a content that we can gesture at by talking about the concept type RABBIT. An instance of this type is a mental symbol that can be used to think about a rabbit as such, or to classify something—perhaps wrongly—as a rabbit; see Fodor. A concept of the type RABBIT-THAT-RAN, which can be used think about something as a rabbit that ran, is presumably a complex mental symbol whose constituents include an instance of RABBIT. A thought can be described as a sentential concept that lets us think about (some portion of) the universe as being a certain way. Thoughts of the type A-RABBIT-RAN can be used to think about the world as being such that a rabbit ran. (4)

As the remarks at the end of this passage indicate, Pietroski takes thoughts to be a *special kind of concept*—namely, a *sentential* concept. This is important to highlight, in view of its relation to a broader point about sentential *meanings*.

Pietroski is skeptical that "Slangs generate sentences as such." The traditional notion of a sentence, as a unity of a subject and a predicate, has been roundly abandoned in contemporary linguistics. While the notions of "subject" and "sentence" have a place in subject-predicate conceptions of *thought*, Pietroski points out that they "may have no stable place" in contemporary scientific grammars (114).

> Linguists have replaced "S" with many phrase-like projections of functional items that include tense and agreement morphemes, along with various complementizers. This raises questions about what sentences are, and whether any grammatical notion corresponds to the notion of a truth-evaluable thought. But theories of grammatical structure—and to that extent, theories of the expressions that Slangs generate—have been improved by *not* positing a special category of sentence. So while such a category often plays a special role in the stipulations regarding invented formal languages, grammatical structure may be independent of any notion of sentence. (61)

Accordingly, Pietroski suspects that talk of "grammatical subjects" is just a roundabout way of "saying that tensed clauses have a 'left edge constituent' that somehow makes them complete sentences—whatever that amounts to—as opposed to mere phrases like 'telephoned Bingley'" (87). Rather than clarifying the notion of a "complete sentence," Pietroski points out that talk of grammatical subjects *presupposes* it.

How, then, to characterize sentences? Naturally, Pietroski does *not* appeal to a distinction between sentential truth conditions and subsentential satisfaction conditions. Instead, he develops a novel version of predicativism, according to which *all* of the concepts assembled by *Begriffsplans* are predicative, in the sense that they all have a classificatory function. This includes concepts that are fetched by linguistic expressions like 'Jessica', 'David Pereplyotchik', and 'Reykjavík'. (Yes, *the* Reykjavík.)

So far, the view on the table is a version of the familiar predicativist position that was introduced by Quine (1970), defended by Burge (1973), and reanimated in contemporary discussions by the work of Delia Graff Fara (2005). Pietroski goes on, however, to make a quite novel claim—namely, that the concepts assembled by *sentence*-sized Slang expressions are *also* predicative.

The idea is that familiar subsentential predicates are assembled, largely via predicate conjunction, and then a new mental operation (⇑ or ⇓) converts the results into a sentential predicate—what Pietroski calls a "polarized concept". Here is how he defines these: "Given any concept M, applying the operation ⇑ yields a polarized concept, ⇑M, that applies to each thing if M applies to something" (30). For instance, if RABBIT applies to something, then ⇑RABBIT applies to each thing and ⇓RABBIT applies to no-thing. We will return to this topic in §3, when we compare this proposal with the inferentialist account of sentence meaning.

Recall that semantic instructions (*Begriffsplans*) have "mechanical execution conditions". Because Pietroski takes *Begriffsplans* to *be* linguistic meanings, it follows for him that that "meanings satisfy demanding compositionality constraints." Such constraints, he argues, permit the assembly of concepts that are better suited for their role in language use than for the epistemic role of "fitting the world". This important upshot of Pietroski's view bears on his rejection of both Davidson's extensional semantics and Lewis's unrestricted type-theoretic approach to natural language (§2.1). For, although he leaves it open that we might build truth-evaluable thoughts as a *side-effect* of language processing, he denies that "meanings are instructions for how to build concepts that exhibit classical semantic properties" (115). Likewise, he suspects that "most natural concepts [do not] have extensions; cp. Travis… if only because of vagueness; cp. Sainsbury" (9). Hence, the *Begriffsplans* that Pietroski identifies with meanings "make no reference to the things we usually think and talk about" (115). If correct, this conclusion is just one more nail in the coffin of the extensionalist project.

2.5. *Pietroski on Fregean thoughts and concepts*

Common to both Pietroski and Brandom is a deep engagement with the work of Frege. However, as we'll see presently, the lessons that Pietroski draws from Frege are *not* those that one might expect. In particular, the formal device that he takes over from Frege's semantics is not that of function application, as is common; rather, he emphasizes Frege's immensely useful notion of concept *invention*—something you don't hear much about in discussions of Frege, at least amongst linguists.

As noted earlier, Pietroski holds that are multiple languages of thought—i.e., distinct formats of concept application. In his discussions of Frege, he advances the hypothesis that there are, in fact, at least two such languages. The first one, in order of evolutionary history, may well have a Fregean semantics and include expressions of type <t>. The second one, which only came in with the evolution of natural language, consists of concepts that were *invented*, or *introduced*, in a Fregean sense, on the basis of the older ones.

> [N]atural sentences of type <t> may belong to languages of thought that are phylogenetically older than Slangs. Expressions of these newer languages may be used to build complex monadic concepts, perhaps including some special cases that are closely related to natural thoughts of type <t>. In which case, the very idea of a truth-conditional semantics for a human language may be fundamentally misguided. (114)

Because Pietroski treats the new type of concept as being invariably predicative—i.e., functioning semantically to classify things into *categories*, not to denote them individually—he calls such concepts "categorical". The older type of concept, which participates in thoughts of type <t>, includes singular denoting concepts and predicates of any adicity. On account of their semantic function of relating items to each other, Pietroski calls such concepts (and the thoughts they participate in) "relational".

Though I see its significance, I'm not, myself, a huge fan of the 'categorical'/'relational' terminology. Adverting to their historical roles, rather than their internal logic, I'll call these languages Olde Mentalese and New Mentalese for the remainder of the discussion. Here's how Pietroski casts the theoretical relations between them.

> Frege assumed that we naturally think and talk in a subject-predicate format, and that we need help—[e.g.] his invented *Begriffsschrift*—in order to use our rudimentary capacities for relational thought in systematic ways… The idea was that a thought content can be "dimly grasped," in some natural way, and then re-presented in a more logically perspicuous format that highlights inferential relations to other contents… I think this is basically right: our categorical thoughts are governed by a natural logic that lets us appreciate certain implication relations among predicates; but our relational concepts are related in less systematic ways. We use relational concepts in natural modes of thought. (95-6)

The distinction between Olde Mentalese and New Mentalese allows Pietroski clarify his perspective, contrasting it with Frege's. Here, too, it's instructive to quote at length.

> Frege introduced higher-order polyadic analogs of monadic concepts. In this respect, my project is the converse of his. Frege invented logically interesting concepts, and he viewed monadicity as a kind of relation to truth, as part of a project in logic that prescinds from many details of human psychology. I think humans naturally use concepts of various adicities to introduce logically boring predicative analogs. But I adopt Frege's idea that available concepts can be used to introduce formally new ones, and that this can be useful for certain derivational purposes. Frege "unpacked" monadic concepts like NUMBER(_), in ways that let him exploit the power of his sophisticated polyadic logic to derive arithmetic axioms from (HP). I am suggesting that Slangs let us use antecedently available concepts— many of which are polyadic—to introduce concepts like CHASE(_) and GIVE(_), which can be combined in simple ways that allow for simple inferences like conjunction reduction. But the big idea, which I am applying to the study of Slangs, is Fregean: languages are not mere tools for expressing available concepts; they can be used to introduce formally new concepts that are useful given certain computational capacities and limitations. This is why I have dwelt so long on Frege's project. For while the idea of concept introduction was important for Frege, it is not the aspect of his work that semanticists typically draw on.

The gory details of Frege's technical devices for concept introduction are, mercifully, beyond our present needs; only a few key points are relevant. One is that introducing concepts need *not* be seen on the model of explicit definition. Rather, Pietroski highlights Frege's proposal for a *second* way of introducing concepts—viz., by *inventing* them. Similarly, although *analyzing* a concept has often been seen as breaking it down into its more basic definitional constituents, Pietroski joins Fodor (1970) in rejecting the idea that lexicalized concepts will generally admit of such analytic definitions. Nevertheless, there is an alternative way of analyzing concepts, which Pietroski characterizes as "a *creative* activity" (emphasis mine).

> Given a very fine-grained notion of content, or thought-equivalence, analysis may not be possible. But Frege employed at least two notions of content: one based on his notion of sense (*Sinn*), and another according to which thoughts are equivalent if each follows from the other. Given the latter notion, or Lewis's characterization of contents as sets of logically possible worlds, one can say that our current representations are not yet perspicuous. We can use our concepts to ask questions that lead us to reformulate the questions in ways that allow for interesting answers. From this perspective, analysis can be a creative activity whose aim is not to depict our current representations…

It's in virtue of our ability to invent new concepts that we, *qua* humans endowed with a specific FL, have invented the monadic and quasi-dyadic concepts that arise only for language use. This includes not only monadic event-predicates like GIVE(_), invented on the basis of the older triadic concept GIVE(*x, y, z*), but also—importantly for Pietroski's pur-

poses, though not ours—*thematic* concepts such as AGENT(_), PATIENT(_), and RECIPIENT(_).

2.6. *Summary*

The generativist methodology that animates Pietroski's inquiry leads him to a number of strikingly original claims about concepts and a detailed theory of meanings. Treating the latter in a resolutely naturalist fashion, he maintains that their theoretical role is to mediate between pronunciations and concepts—i.e., to effect the psychological operations that constitute the interface between language (FL) and the "conceptual-intentional system" (to use Chomsky's coinage). Although meanings facilitate the *assembly* of concepts, which *have* intentional contents, Pietroski holds that meanings are neither concepts *nor* their contents.

On this view, the relation between truth and conceptual/intentional content is "quite complicated and orthogonal to the central issues concerning how meanings compose" (115). This, among the many other reasons surveyed above, leads Pietroski to abandon Davidson's project of extensional truth-conditional semantics. Moreover, the goal of *explaining* our access to a productive hierarchy of concepts, rather than merely *stipulating* it, underlies his rejection of the type-theoretic approach championed by Lewis (1970)—one of the many disagreements that we'll look at in the next section.

The semantic theory that satisfies Pietroski's methodological commitments—as well as the compositionality constraints that he argues follow from it—treats meanings as composable instructions for concept assembly. The instructions are "composable" in the sense that their basic constituents—namely, *fetch@* and *join[I, I']*—can enter into part-whole relations to one another. Moreover, as noted earlier, the larger structures they compose will, in a definite sense, *mirror* those of the concepts that the instructions assemble.

Having furnished empirical evidence for the idea that these "*Begriffsplans*" reduce largely to two flavors of *predicate conjunction*, Pietroski adopts a strong version of predicativism, according to which *all* of the concepts that natural language allows us to access and assemble are predicative. This includes not only the concepts fetched by linguistic expressions that have *traditionally* been classed as predicates, but also those that have generally been seen as differing in some important respect—including singular terms and, more strikingly, even *sentences*. The conceptual predicates that meanings allow us to access and assemble thus all either monadic, dyadic (in a restricted sense), or "polarized", where the latter kind is assembled by sentence-like linguistic expressions, using specialized mental operations, ⇑ and ⇓, to "polarize" concepts. Importantly, the resulting conceptual structures are not necessarily ones that best "fit the world", and they're not even the only ones we can deploy in thought. But, if Pietroski is correct, the they *are* the only ones that FL can assemble.

Denying that the concepts involved in language use have denotational properties and relational structures (of arbitrary adicity) leaves open whether *other* concepts might have these features. As we saw, Pietroski hypothesizes that there *are* in fact such concepts, and that they belong to a phylogenetically older language of thought than the one FL allows us to access—what I've dubbed 'Olde Mentalese'. Olde thoughts might have a subject-predicate form, a Fregean semantics, and belong to the semantic type <t>.

Pietroski goes on to make novel use of Frege's notion of concept *invention* in explaining the (non-definitional) mental introduction of *new* concepts on the basis of the Olde ones—specifically, the ones that FL allows us to access/assemble (New Mentalese). This psychological process, he argues, serves to introduce GIVE(_) on the basis of GIVE(*x, y, z*), as well as novel *thematic* concepts such as AGENT(_), PATIENT(_). These, in turn, participate in building polarized sentential concepts, such as ⇑RABBIT, which "applies to each thing if RABBIT applies to something". In the course of assembling such concepts, it may happen—*but only as a side-effect* (fortuitous or otherwise)—that we *also* token thoughts of Olde Mentalese. But the details of how Olde Mentalese thoughts functions are, Pietroski rightly holds, beyond the scope of a naturalistic semantic inquiry into human *language*.

Conclusion of Part 1

We've now surveyed the core commitments of two large-scale theoretical frameworks in the philosophy of language and seen some of the ways in which they play out in the realm of semantics, including in detailed analyses of various linguistic constructions. It may appear that the two views are so different in substance and overall methodology that a conversation between the two is unlikely to bear much fruit. In fact, I suspect this is a large part of why so few conversations of this kind ever take place. In Part 2 of this essay (next issue), I'll argue for a contrary perspective, outlining an ecumenical approach that seeks to integrate the two in a variety of ways. In surveying what I take to be significant points of convergence—which then serve as background for constraining residual disputes—I rebuff various superficial objections to the possibility of integration. In each case, I show how the theoretical differences that they point to can be reconciled without doing much (if any) violence to either view.

References

Brandom, R. B. 1987. "Singular Terms and Sentential Sign Designs." *Philosophical Topics* 15 (1): 145–167.

Brandom, R. B. 1994. *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge: Harvard University Press.

Brandom, R. B. 2001. *Articulating Reasons: An Introduction to Inferentialism*. Cambridge: Harvard University Press.

Brandom, R. B. 2008. *Between Saying and Doing: Towards an Analytic Pragmatism*. Cambridge: Harvard University Press.

Brandom, R. B. 2009. "Why Philosophy Has Failed Cognitive Science." In *Reason and Philosophy: Animating Ideas*. Cambridge: Harvard University Press.

Burge, T. 1973. "Reference and proper names." *Journal of Philosophy* 70: 425–439.

Burge, T. 2010. *Origins of Objectivity*. New York: Oxford University Press

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

Chomsky, N. 1986. *Knowledge of Language*. New York: Praeger.

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

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

Chomsky, N. 1995. *Why Only Us: Language and Evolution*. Cambridge: MIT Press.

Collins, J. 2012. *The Unity of Linguistic Meaning*. Oxford: Oxford University Press.

Davidson, D. 1983. *Essays on Truth and Interpretation*. Oxford: Oxford University Press.

Dummett, M. 1978. *Truth and other Enigmas*. London: Duckworth.

Fodor, J. A. 1975. *The Language of Thought*. Oxford: Oxford University Press.

Fodor, J. A. 1979. "Three Reasons for not Deriving 'Kill' from 'Cause to Die'." *Linguistic Inquiry* 1: 429-438.

Fodor, J. A., Bever, T. and Garrett, M. 1974. *The Psychology of Language*, Cambridge: MIT Press.

Fodor, J. A. and Lepore, E. 1992. *Holism: A Shopper's Guide*. Oxford: Blackwell.

Fodor, J. A. and Pylyshyn, Z. W. 2015. *Minds Without Meanings: An Essay on the Content of Concepts*. Cambridge: MIT Press.

Grover, D. L., Camp, J. L., Belnap, N. D. 1975. "A Prosententialist Theory of Truth." *Philosophical Studies* 27 (2): 73–125.

Graff Fara, D. 2015. "Names are Predicates." *Philosophical Review* 124 (1): 59–117.

Grice, H. P. 1989. *Studies in the Way of Words*. Cambridge: Harvard University Press.

Harris, D. W. 2020. "We Talk to People, not Contexts" *Philosophical Studies* 177 (9): 2713–2733.

Heim, I. and Kratzer, A. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell.

Horwich, P. 1990. *Truth*. Oxford: Oxford University Press.

Kripke, S. 1980. *Naming and Necessity*. Cambridge: Harvard University Press.

Levinson, S. C. 1983. *Pragmatics*. Cambridge: Cambridge University Press

Lewis, D. K. 1969. *Convention*. Cambridge: Harvard University Press.

Lewis, D. K. 1970. "General Semantics." *Synthese* 22: 18–67.

Lewis, D. K. 1973. "Languages and Language." In K. Gunderson (ed.). *Minnesota Studies in the Philosophy of Science*. Vol. 7. Minneapolis: University of Minnesota Press.

Pereplyotchik, D. 2017. *Psychosyntax: The Nature of Grammar and Its Place in the Mind*. Springer International.

Pereplyotchik, D. 2019. "Whither Extensions." *Mind and Language* (Special Issue) 1–14.

Pietroski, P. M. 2018. *Conjoining Meanings: Semantics Without Truth Values*. Oxford: Oxford University Press.

Price, H. 2011. *Naturalism Without Mirrors*. Oxford: Oxford University Press.

Price, H. 2013. *Expressivism, Pragmatism and Representationalism*. Cambridge: Cambridge University Press.

Quine, W. V. 1953. "On What There Is." In *From a Logical Point of View*. New York: Harper and Row.

Quine, W. V. 1960. *Word and Object*. Cambridge: MIT Press.

Quine, W. V. 1970/1980. *Philosophy of Logic, 2nd Edition*. Cambridge: Harvard University Press.

Rorty, R. 1979. *Philosophy and the Mirror of Nature*. Princeton: Princeton University Press.

Sellars, W. 1963. "Philosophy and the Scientific Image of Man." In R. Colodny (ed.). *Science, Perception, and Reality*. Humanities Press/Ridgeview: 35–78.

Stainton, R. J. 2006. *Words and Thoughts: Subsentences, Ellipsis, and the Philosophy of Language*. Oxford: Oxford University Press

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

Sperber, D. and Wilson, D. 1986/1995. *Relevance*: *Communication and Cognition*, *2nd Edition*. Oxford: Blackwell.

Tomasello, M. 2005. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge: Harvard University Press.

Yang, C. 2006. *The Infinite Gift: How Children Learn and Unlearn the Languages of the World*. New York: Scribner.