Mind & Language, ISSN: 0268-1064
Vol. 13. No. 2 June 1998, pp 225–280.

Meanings Are Syntactically Individuated and Found in the Head

JAMES MCGILVRAY

Abstract: Expanding on some of Chomsky's recently expressed views of meaning in a way that is consistent with his long-held rationalist conception of mind, I show how syntax, broadly conceived, could individuate meanings and provide a science of meanings inside the head. Interpretation becomes a pragmatic matter, although a rationalist account of mind shows how internal meanings guide interpretation and, more generally, language use. In this view of meanings, interpretation, and mind, semantics as usually understood disappears.

1. Introduction: Expressions in the Head

Most philosophers and linguists look for a theory of the meaning or semantic content of natural language expressions in the relationship of linguistic expressions to the world—they look to truth, reference, belief, language games, intentionality, and the like. Call this the *externalist* approach. I explain here why one should look instead to expressions inside the head for meanings; this is the *internalist* approach. The basic thesis is that meanings are contents intrinsic to expressions (intrinsic contents), and that they are defined and individuated by syntax, broadly conceived. Terminological differences aside, this is basically Chomsky's view as it appears within the minimalist form of the principles-and-parameters framework. I outline the semantic claims and implications of the minimalist program below. The basic strategy of this program—lexical/morphological feature-elimination aiming towards convergence at the phonetic interface PHON (P[honetic]F[orm]) and

I gratefully acknowledge research support from an FCAR équipe grant. Parts of an early draft of this paper were presented at the Canadian Philosophical Association meetings in Calgary in June, 1994, and other parts at the 1995 CPA meetings. I want to thank an anonymous reviewer for *Mind & Language* and Sarah Patterson for suggestions and help. And I particularly want to thank Mark Baker, Noam Chomsky, David Davies, Sue Dwyer, Melinda Hogan, Norbert Hornstein, Ann McKenzie, Claude Panaccio, and Paul Pietroski for very helpful comments and criticism of earlier work. While each helped a great deal, none of them is—of course—to be blamed for the result.

Address for correspondence: Department of Philosophy, McGill University, Leacock Building, 855 Sherbrooke Street West, Montreal, Quebec, Canada H3A 2T7.
Email: jim@philo.mcgill.ca.

the semantic interface SEM (L[ogical]F[orm])—is consistent with and advances Chomsky's complex of internalist and nativist semantic views, views that he has maintained in their basic form for about forty years. It also provides the framework for an honest internalist science of meanings, or intrinsic contents.

This unusual approach to meaning has few contemporary supporters. Chomsky cites some whose works contribute to various aspects of his recent efforts to develop an internalist, nativist science of meaning (e.g. Moravcsik, 1975, 1990, Pustejovsky, 1995), but to find anyone who not only thinks that meanings are in the head but who advocates the nativist presuppositions and pursues the constructivist implications of this view, one must go back to philosopher-scientists in the tradition of 'Cartesian linguistics' (Chomsky, 1966). For example, two Cambridge Platonists in that tradition, Ralph Cudworth and Herbert of Cherbury, not only held that meanings (innately fixed 'ideas' or 'notions') are in the head, but they try to speak to how these meanings in the head bear on the world.[1] As Cudworth says (1731, p. 133): 'individual things existing without the soul, are but the secondary objects of knowledge and intellection, which the mind understands not by looking out from it self as sense doth, but by reflecting inwardly upon it self, and comprehending them under those intelligible ideas or reasonings of its own, which it protrudes from within it self; so that the mind or intellect may well be called the measure of all things'. Cudworth had little inkling of the generative computational theory that could provide a genuine science of meanings or 'ideas', of course, but his 'protrudes' can still be read as 'generates'. The basic idea behind a generative system was bruited about at the end of the sixteenth century by Huarte and others, and it was not uncommon to marvel at the way in which such a small number of letters could produce all the words in a language. In any case, with that science in place and Cudworth's 'ideas' transformed into Chomsky's linguistic interfaces (SEMs) and his 'individual things existing without the soul' made into the objects of the common-sense world, we can have the mind 'protruding' ideas to 'make' the things of perception and common sense. The result is an internal-

[1] Cudworth, posthumous 1731, reprint 1976); Herbert of Cherbury, 1937. Chomsky, 1966, cites both works and the first in Chomsky, 1997. Herbert in 1622 tells us of our sorry cognitive state if we did not have the guidance of internal, innate ideas: 'Vacant forms, prodigies, and fearful images would pass meaninglessly and even dangerously before our minds, unless there existed within us, in the shape of notions imprinted in the mind, that analogous faculty by which we distinguish good from evil. From where else could we have received knowledge? In consequence, anyone who considers to what extent objects in their external relationship contribute to their current perception; who seeks to estimate what is contributed by us, or to discover what is due to alien or accidental sources, or again to innate influences, or to factors arising from nature, will be led to refer to these principles. We listen to the voice of nature not only in our choice between what is good and evil, beneficial and harmful, but also in that external correspondence by which we distinguish truth from falsehood, we possess hidden faculties which when stimulated by objects quickly respond to them' p. 106.

ist, nativist and constructivist cognitive framework that updates Cudworth's view that the mind is, in a sense, 'the measure of . . . things'. As Chomsky puts it (1966, p. 65), 'The strong assumptions about innate mental structure made by rationalistic psychology and philosophy of mind eliminated the necessity for any sharp distinction between a theory of perception and a theory of learning'. What the mind can perceive and understand depends on what it brings to perception and understanding, and does not come 'from without'. Certainly this seems clear from contemporary studies of phonological/phonetic capacities (Mehler *et al.* 1994); perhaps we should think in the same way of perception of events such as givings and things like water and cities. If so, an internalist account of meanings suits only a constructivist account of 'interpretation'—of how inner meanings bear on the description and explanation of things.

One contemporary Chomsky does not often cite approvingly (although see Chomsky, 1997), but who is useful for the idea that meaning is in syntax as well as for getting a grip on Cudworth's constructivist suggestion that the mind 'protrudes' ideas that are used to make things and events, is Nelson Goodman. Goodman (1952, 1961, 1968) points to various ways in which to classify a linguistic expression in terms of what he calls its 'syntax' (that is, independently of what the expression denotes or refers to), where this form of classification yields a kind of content intrinsic to the syntactic entity. For example, the phrase 'under the spreading chestnut tree' would be classified by Goodman (1968) as an under-the-spreading-chestnut-tree kind of phrase (here, a prepositional phrase); syntactic classification of this sort yields a kind of content because being able to classify the phrase in this way, thereby distinguishing it from other phrases, just *is* to know its content. On its own, Goodman's suggestion does not appear promising. In fact, if it is based on his crude concept of syntax and his externalist and anti-nativist view of how expressions come by their 'syntactic' classifications, it is doomed. Perhaps this is the reason that Chomsky goes to the seventeenth and eighteenth centuries to find more plausible presuppositions. But if Goodman's suggestion is placed within Chomsky's sophisticated computational syntactic theory—one that allows for extremely fine-grained individuation of all the features found in the expression—and is fleshed out by providing a nativist cognitive framework that shows how an expression (syntactically characterized entity that it is) offers a language user in its semantic interface an internally constituted 'perspective' that can play a role in individuating and even constructing the things of a world, it becomes not only plausible, but compelling. The end result is a conceptualist-nominalist view of the meanings of linguistic expressions that participate in what Goodman called 'worldmaking', a view that—if he could overcome his anti-nativism—Goodman would welcome. In this view, meanings of linguistic expressions become linguistic entities in the head and are treated in the way a conceptualist and constructivist wants; these concepts provide perspectives that the speaker can use for any number of purposes, including referring to and describing the

objects of a world, thereby (as Goodman puts it) 'making' a world.[2] Crucially, however—and this is very unlike Goodman—these concepts are individuated by internally and innately specified features, not by their relationships to the world, if any. They are, I suggest, identical with Chomsky's semantic interfaces—SEMs or LFs.[3]

To see where this leads, it is worth asking what happens to semantics as it is more-or-less standardly conceived (that is, as an account of the relationship between language and world) after the internalist view is adopted. Adopting it suggests partitioning externalist approaches and the phenomena they investigate. Some, or at least parts of some approaches and phenomena, can be appropriated by the (broad) syntactician and placed within syntax. I explain what this involves by discussing a paper of Larson and Ludlow, two friends of LFs, in section 4. The rest—any that explicitly or implicitly take their task to be solely that of dealing with how natural language expressions are used by persons to effect various kinds of ends (including correctly describing their world(s))—are consigned to a pragmatic domain where, while they might be able to contribute data to the syntactician, for the most part they can hope at best to provide insightful descriptions, not of linguistic expressions and their meanings, but of how these expressions with their meanings are used. Think here of Charles Morris's tripartite division of approaches to signs in terms of syntax, semantics, and pragmatics. Syntax deals with sign-to-sign relations, semantics with relations between signs and the world or 'interpretation', and pragmatics with the complex relations that arise in the use of signs by speakers. Replacing Morris's sign-to-sign relations with a sophisticated morphosyntactic and internalist version of generative grammar, one can say that Chomsky proposes eliminating semantics by moving an account of content into syntax, broadly conceived, and by relegating language–world relations or interpretation to the domain of language use, or pragmatics. Semantics disappears as a separate field, and what remains is (broad) syntax, of which there is a developing science, and pragmatics, where one can really only hope for perspicuous description, sometimes systematic. On the pragmatic side, Chomsky seems to agree with Wittgenstein of the *Investigations*: there is no theory of the domain of language use, just more or less helpful description, systematic only because it is

---

[2] Not all concepts are linguistic, for our cognitive capacities classify in various ways (e.g. in terms of colours, shapes, and sounds). Here, I focus on linguistic concepts. They are specific to the language faculty, not to be found elsewhere.

[3] LFs or SEMs are syntactically specified, period. While differences in SEMs make for differences in interpretations (or are likely to), interpretational differences—if this means differences in truth values, referents, or the like—are irrelevant to the individuation of SEMs. What counts in their individuation is just being syntactically well-computed—not 'crashing', in the minimalist framework. This principle is not always respected by some semanticians who are otherwise friendly to LFs: Higginbotham, 1985; May, 1985; Larson and Ludlow, 1993.

based on human projects, or tasks.[4] On the syntactic side, he offers a developed scientific theory. Fortunately for my purposes—and those of anyone else who wants a serious theory—meaning, or SEM, is within the domain of this science.

A caveat: while partitioning semantics into (broad) syntax and pragmatics is good internalist strategy, this is not the whole story. Chomsky insists that one can speak as a scientist to *some* aspects of language use, so long as these are internal to the head. Thus, the language faculty, through PHON and SEM, provides 'instructions' to other internal systems—in the case of PHON, to speech production and audition systems. And as I hinted above, it is possible to provide a better, nativist picture of what is involved in language use and 'interpretation'. One can do so by speaking to how internal and largely innately fixed features of expressions 'guide' their use in developing and providing knowledge of the world. This adaptation and updating of Cudworth and Herbert's picture of the innately fixed mind and its forays into the world provides more insight than simple description. It also helps show why the use of an expression with the *meaning* that it has is also plausibly *meaningful* to us, why it contributes to the fulfilment of some task or another.[5]

---

[4] This is related to the point made by Davidson, 1986, in a way that might become clear below in section 4. On some interpretations of 'Derangements' Davidson is denying that there can be a theory of language, period. This makes Davidson easy to refute: obviously, there are syntactic theories of language, and they are honest theories. A more interesting claim is this: in spite of his earlier efforts to produce a truth based theory of meaning (or interpretation), Davidson is just denying that there can be a theory of meaning for a language, where 'meaning' is understood in his, not Chomsky's way. I suspect that this is Davidson's view. Those taken with the Davidsonian project of constructing an 'innocent' semantics try, however, to avoid collapsing interpretation into what looks like haphazard pragmatics. Pietroski, 1994, argues that Davidson's paper is not an outright rejection of theories of interpretation, but in fact a defence of the idea that a proper theory of interpretation based on truth is (in principle) available. For Davidson, on Pietroski's view, the condition for getting a proper theory in the domain of the relation of language to world (interpretation) lies in excising *ad hoc* 'passing theories'; what is left is—it is suggested—a proper theory, based on knowledge of truth conditions. Unlike Pietroski, I deny that 'what is left' after excising what Davidson calls passing theories is a theory. Something may indeed be left: pragmatics is not all Malaprop cases. If one limits one's observations about uses of language to finding truths-in-a-community or some such, one can find a reasonably systematic set of judgments which as a core set of beliefs can serve as a basis of translation and interpretation. One gets an idealized version of 'what we think', where 'we' is a selection of some *cognoscenti* for some enterprise. But however systematic—approximating in some cases perhaps to a 'logic' for an enterprise—it is not a theory, but a normatively governed idealization of usage. On the other hand, one *can* have an honest theory of meaning by going to syntax.

[5] I use the terminology suggested here in McGilvray, 1991, a pre-minimalist (and slightly idiosyncratic) effort to deal with the implications for tense of a division between the meaning of an expression and its use. The term 'meaningful' recalls the pragmatic tradition with meaning, e.g. C.I. Lewis. I suspect that an account of what makes an expression meaningful is what most people seek when they ask for 'the meaning of an expression' —an account, or demonstration, of its use by someone to fulfil some task or another. This also suits the idea that meaning has something to do with reference and truth—with relating expressions to the world.

A policy: I use underscoring as a device to orthographically represent SEMs or LFs, as in under the tree (a P[repositional] P[hrase] SEM) She quickly strode into the room (a S[entence] SEM) and miss the target still again (a V[erb] P[hrase] SEM). Treat underscoring as an invitation to take the expression that the orthographic marks on a page prompt in your head and use that prompted expression to refer to its SEM—its own meaning. I explain later that this is a form of exemplificational reference. If the invitation is accepted, and the referential task performed, according to the thesis defended here one refers—although not necessarily under the following description—to a set of features that individuate the meaning (SEM) of that expression. Complementary, more formal ways to refer to SEMs include labelled strings and phrase structure diagrams. For example, one could have $[_{PP}[_{P}$under$]$ $[_{DP}[_{D}$the$]$ $[_{N}$tree$]]]$ (note the continued underscoring), or its equivalent phrase structure diagram. Alternative names for SEMs include 'property' or 'relation' (for VP, PP, and A[djective] P[hrase] SEMs), 'intension' (for N[oun] P[hrase] SEMs), 'situation' (for S, C[omplement] or I[nflectional] P[hrase] SEMs), and the all-purpose 'concept'. Abstract entities like properties, relations, and situations plus non-syntactic mentalist entities like concepts, are deflated into syntactically individuated SEMs; I explain this in section 3. Clearly, syntax cuts a wide swath in this account: meanings are placed in expressions which are in turn placed in the head, providing an exciting prospect for the nominalist.

2. Computation and Content

A serious theory of meaning cannot be what sometimes passes informally for theory among philosophers and others—anything from a conjecture to a set of reasonably systematic claims about, say, beliefs and knowledge and their role in behaviour. Nor can it be merely an elegant but misguided formalism that deals with the wrong domain and/or has no explanatory warrant. I will argue that to succeed, a theory of meaning for a natural language must be part of a syntactic theory. It must then *de facto* be a *computational* theory that describes and relates internal (mental) states and processes and explains the language acquisition data represented in poverty of stimulus observations. Specifically, it must explain why and how children manage to acquire the meanings of natural language expressions so readily while relying on very limited evidence. If it succeeds at this, a theory of meaning becomes a contribution to the universal theory of syntax—universal across the human species. Chomsky's view—now usually accepted, albeit often grudgingly and with many reservations—is that only a nativist or innatist and biological theory can explain poverty of stimulus observations. This nativist approach takes a particularly interesting turn in the case of a theory of meaning, for a nativist must accept that a great deal of a person's lexicon is innate—something even those who sympathize with Chomsky about the innateness of phonological and structural features of language find difficult

to accept. I take this up in the next section. Here I try to deal with some philosophical problems that arise when one takes a computational theory and suggests having it deal with intrinsic contents—with meanings in the head.

Many attempts to provide a theory of human cognitive competences, such as vision, audition, and taste, are computational.[6] If computational, a theory is—adapting Fodor's (1980) terminology—a *formal* theory. To be formal in this sense is not necessarily to deal with structural features alone; a theory can be formal, and yet deal with the fine-grained features to be found in its domain. Rather, this formality condition insists that the theory introduce, define, and relate by computationally tractable algorithms symbols that deal with the states/processes/events of its domain. In the case of Marr's computational theory of vision, for example, to be formal is to be mathematical: inputs to vision are response intensities that in the computational theory are specified numerically, so that the formal symbols of the theory are numbers, and a computational algorithm like that found in 'edge detection' takes arrayed numbers to solve for zero crossings ('edges'). In the case of Chomsky's computational theory of linguistic sound and linguistic meaning in its present, minimalist form, to be formal and computational is to be syntactic-morphological. Computations in the head itself ('language processing') proceed locally over the *features* that make up lexical items. This processing is largely an elimination procedure, taking as input sets of lexical-morphological features of three sorts (phonological, formal, and semantic) and converging to two separate feature arrays called PHON or PF (sound) and SEM or LF (meaning), where they must meet 'bare output conditions' imposed by the systems with which they interface. The symbols for the theory that deals with these computations are, then, feature-denoting syntactic-morphological terms like 'N' and 'ABSTRACT', and the theory relates these by algorithms like Merge. Some details appear below.

If formal, a computational theory of a cognitive competence need not be restricted to what is in the head. But because computations must also be local—they must proceed over 'neighbouring' elements, or be local—practical necessity restricts it to items in the head. Extending the domain of a theory of a cognitive competence outside the head makes it responsible for far too much; a theory would have to cover so much that it would lose any serious chance of being constructible. This poses a problem for an intentionalist view of a cognitive capacity, one that insists on determining internal states by their relations to things outside the head. Intentional properties arise when an inner state or process is treated as a representation of something in the world, or as having referential or alethic properties. That 'John'

[6] Perhaps these computational theories will be conjoined with (neuro) physiological accounts some day, but as it stands, no neurophysiological account can do more than gesture towards the task of individuating colours, sounds, shapes, and—pre-eminently—linguistic expressions, much less show how the multiplicity of combinations of these are actually produced.

is used by Mary on an occasion to refer to John is an intentional property of 'John' on that occasion. Intentional properties like this cannot, I think, play any *essential* role in the individuation of inner cognitive states, nor the formal symbols of the theory: that is, they cannot be allowed to play any role in the computational theory. Because they insist on relating a computational theory's domain to users of cognitive states and to things outside the head, they expand the domain of the theory and make it unmanageable. Because they require that computation respect demands that cannot be determined locally, they undermine any serious chance of constructing a computational theory of intentionally specified cognitive states at all—or so I argue.

At this point many readers, if they have not already, will rebel, exclaiming that nothing semantic like meanings can be found in the head alone. Furthermore, surely nothing syntactic can be semantic. And if that were not enough, the nature of a computational theory has been woefully misconstrued. Why, Marr explicitly made computational theories intentional: he said—it is claimed—that computational theories *are* theories of how an organism relates to the world, so that computational states (the relevant symbols) are essentially intentional. A computational theory of a cognitive competence is, in effect, essentially externalist and extrinsic. However difficult that may make it to construct, then, I am wrong to make a computational theory of a cognitive competence internalist and intrinsic.

While the attitude in this set of responses has seemed to many to be persuasive, and while it pervades discussions of the relationship of syntax to semantics, it is this attitude and the externalist claims that go with it that are wrong. To see why, let us begin with the belief that Marr held that a computational theory of a cognitive competence is essentially intentional. Showing that Marr did not hold this is more difficult than it might be, because Marr himself does not help. He sometimes explicates his views in Gibsonian terms, and Gibson held an intentionalist view of cognitive states, if anyone did. Furthermore, Marr speaks of computational states as 'representations'. Now, Marr's Gibsonian remarks can, I think, be dismissed as informal explication, perhaps misguided, but clearly not an essential part of his computational theory.[7] Using the term 'representation', however, invites

[7] There are tradeoffs here. Marr chose to point his readers' understanding of the cognitive implications of his computational theory in certain directions by appealing to a readily understood perceptual psychologist and—I believe—he misled in doing so. Chomsky, in contrast, chose from the start to point his readers' understanding in certain directions by appealing to a little-understood, in fact often systematically misunderstood, tradition in the history of philosophy, science, and linguistics: he appeals to Descartes, to Arnauld and Cordemoy, to Cudworth and the Cambridge Platonists, and to Schlegel and Humboldt. And he has pulled no punches in telling most of his contemporaries in linguistics, psychology, philosophy, and now cognitive science that, with few exceptions, their work is misconceived—not only not an improvement on work done three centuries ago, but a regression. It is not surprising that his views on the place of language in cognition are neither easily understood nor liked. But—I suspect—what he loses in taking the high road he gains in persuading us to rethink the nature of the task of constructing a theory of language.

misunderstanding. The word 'representation' suggests to many philosophical ears—in some contexts, correctly—something like this: whatever else a representation is, it is essentially something that re-presents the things of the world. Hence, a representation is essentially intentional. However, Marr's theory in no way requires that what *he* called representations be individuated in terms of their relations to things in the world. When he spoke of representations, he spoke only of intensities and the like over which computations proceed; these are certainly not *essentially* intentional. And there is also an easy way to avoid being misled: give up the term 'representation' when speaking of the subject matter of a computational theory, and speak of intensities denoted by numbers (Marr) or syntactic-morphological features denoted by appropriate feature terms (Chomsky).

If these factors can be dismissed, another cannot. The core of the externalist, intentionalist interpretation of Marr lies in what he says about the nature of a computational theory. He held that one must choose one's algorithms to serve computational functions. This, then, is the issue: could he, or anyone, have thought that vision's function is anything *except* enabling an organism to get around in its environment? If there is no alternative, the intentionalist must be right. But there is. While there is no doubt that it is important to decide what function the computational principles and symbols of a theory are to serve before choosing the algorithms and formal symbols that deal with these functions, the only obvious function they need serve in Marr's theory is to detail how the visual system operates inside the head. Marr's claim about having to decide upon function before proceeding becomes, then, the commonplace that one must decide what domain the formal symbols and algorithms are supposed to deal with before one can proceed to construct the theory. This amounts to saying that one must decide what kind of theory it is—of visual processing? auditory processing? language processing? Once this is settled, one can begin to choose appropriate formal symbols and algorithms for the relevant domain. It would not do, for example, to choose formal symbols like 'D[eterminer]' and 'NP', operations like Merge and Move/Attract, and the principles of 'binding theory' to deal with vision; for vision, Marr naturally and correctly chooses numerical values and Gaussian functions. There is nothing in *this* view of the function that a computational theory must serve, nor in the choices of algorithms it determines, to push the domain of the theory outside the head. It is, in fact, a theory of what happens in the head.

It helps in keeping a grip on this point, I think, to borrow 'faculty' terminology and ask what function the vision faculty has that is essential to its being the faculty that it is. It is enough, surely, to say that its function is visual processing; this is all that Marr's claim about function need insist on. Of course, one can add that an organism typically employs or uses the output of this faculty to deal with its world, and in doing so gives an output of the vision faculty the *supplementary* intentional function or role *for the organism at a time* of dealing with some object or objects in world. And there is no doubt too that human vision would not have developed into the faculty

it is over some 25 million years (assuming that this faculty is relevantly like those of other primates) unless it routinely and successfully did get so used. But neither of these considerations require that the function of the visual faculty that figures in constructing a computational theory of vision be defined in terms of the use to which it is put by an organism, nor—*a fortiori*—that the history of this use by this organism and phylogenetically related ones need be taken into account.

It also helps to see this point if one asks whether Marr's theory of vision is necessarily a computational theory of human *behaviour*. If it were, it would have to provide formal symbols to capture all relevant stages of the procedure of visual information-gathering and visually based organismic activity, and would have to relate these by various sorts of appropriate algorithms. Clearly, however, Marr's theory with its algorithms deals only with a restricted set of processes in the head. Everything else is ignored—the activities of photons at various energy levels, the variable causes of reflection and diffraction, the roles of media, the processes involved in memory, motor processes of an indeterminate number of kinds, the processes involved in affecting things . . . As I suggested, restricting the domain of the theory to a restricted class of happenings in the head is one key to success—here and in other cognitive domains. To extend the domain of a computational theory of a cognitive competence to things outside the head, and within the head to other cognitive domains, is to broaden the task of the theory to the point that it becomes unmanageable: it undermines the effort to construct an honest theory. This is true of vision.[8] And it is particularly obvious with computational theories of language, for in the case of language one must also take into account what Chomsky calls the 'creative aspect of language use'. Language use is *stimulus-free*, *unbounded*, and yet, in use, typically *appropriate* to any number of tasks. Because of this, 'causal' (informational, etc.) accounts that link words to circumstances outside the head and hope to make these links essential to the word are hopeless. Moreover, there is no hope for a 'normal' function for language; not even human vision is 'devoted', but language is certainly not. Prospects for an honest theory of meaning that insists

[8] Something reasonably close to a computational theory of some aspects of vision, such as colour vision, can be pushed 'outside', but at heavy cost. To do this, one must make very strong assumptions about a large number of factors in organism and environment; I take this up briefly in section 4 in connection with a discussion of whether there can be a computational theory of language interpretation. But even if such a move is made and it turns out to be moderately successful, it should be called a computational theory of colour vision *plus media and the distal causes of colour vision*. The theory of *colour vision*, and the theory of *colour* itself, is and remains inside the head alone: nothing in the world 'outside' answers to the structure and texture of colours. Thus, the basic claim that a computational theory of a cognitive capacity is restricted to what is in the head remains correct; the theory of colour vision is through-and-through internal. (See McGilvray, 1994). In any case, not even this tentative way of taking computations outside the head works with language. As the creative aspect indicates, language is not a 'devoted' capacity in the way colour vision can—at a cost—be thought to be.

on individuating meanings intentionally are thus nil.[9] Those looking for a serious theory of meaning should, then, look inside the head.

Another point—related to the fact that language use is stimulus-free—reinforces my claim that a computational theory should be internal, and restrict itself to locally determined features; this argument shows that computational states ('symbols'—here, numerical values with vision and morphological-lexical features with language) cannot have their intentional features essentially. The argument is simple. If they were to have them essentially, the organism's environment would become essential to their individuation. But, obviously, it is not. As Egan (1995) points out, if an organism with a functioning computational system is not in its normal environment, however that is construed, while the algorithms in its computational systems will still serve their computationally (mathematically and internally) specified functions to the extent that there are processes of the relevant types at all, they will not serve their normal intentional functions (assuming that they have a normal intentional function—which language does not). So, clearly, they do not have their intentional contents essentially.

So the externalist attitude and the claims that cluster around it seem to be wrong. If there is to be a computational *theory* of meaning that properly individuates meanings by defining formal symbols and shows how to generate complex meanings by introducing algorithms that manipulate these symbols, nothing says that this theory need deal with anything outside the head, practical considerations lead one to avoid going outside the head, no serious theory actually does go outside the head, and the creative aspect of language use suggests that going outside the head is disastrous. Of course, as with vision, there is no reason to deny that these syntactically specified processes that have (are) meanings intrinsically can also gain other properties when used to perform tasks on occasions—such properties as being used by a person to refer to a wombat or a widget, perhaps making irascible, loquacious, silly, or even true, judgements about that wombat or widget. But these other properties are by no means essential to meanings, or intrinsic contents (SEMs/LFs). Nor is dealing with these addenda to formally defined meanings inside the head—intrinsic contents—within the domain of a theory of meaning,[10] although we can usually *describe* how a person uses SEMs,

---

[9] If the intentionalist thinks that 'symbols' get intentional content by being causally related to the thing(s) that the 'symbol' has within its ('normal') extension, it may be enough to be stimulus free.

[10] Methodologically, of course, in constructing the theory, one has to appeal to observable phenomena and in that sense to the external world. But that does not mean that the formal classification that one ends up with is in any way essentially distal. For discussion of a parallel case concerning colours, see Clark, 1993, and my further discussion of Clark's claims, in McGilvray, 1994. I do not, by the way, claim that one can construct a theory of language in ways that have proven successful in constructing a theory of sensory states, such as colours. I seriously doubt, for example, that multi-dimensional scaling can tell us anything about the syntax of a language. There are no 'dimensions' here in the way there are with colours.

given the tasks that s/he is performing, and a great deal more contextual information.

To ward off possible misunderstanding, intrinsic contents are not narrow contents. First, narrow content is often assumed to be defined over the physical states of an individual, where intrinsic contents have no such pretensions.[11] (No attempts to actually define narrow content in this way (e.g. Block, 1986; Segal, 1991, have proven convincing, although that is not the point.) Second, narrow content is essentially intentional: it is meant to explain an individual's actions in those cases where s/he is outside of a 'normal environment'. Intrinsic contents (SEMs) are not introduced to explain behaviour. Of course, if you want to appropriate the term 'narrow content' and use it for intrinsic contents, you are probably free to do so (see Chomsky, 1997). A technical term due to Fodor, he seems to have virtually abandoned it (Fodor, 1994).

3. Meanings in the Head

3.1 Computation and Creativity

I will assume in what follows that a computational theory of meanings must provide for a way to describe the meanings of 'words', or lexical items; it must provide a principled way to distinguish one meaning from another, to individuate meanings. For Chomsky, this is done with feature-denoting theoretical terms. It must also—and this is a task to which a computational theory is particularly well suited—show by appeal to principles how the meanings of individual lexical items combine to produce complex meanings. And finally, it must speak to the poverty-of-stimulus facts; in effect, it must make meanings innate.

It is important not to underestimate the extent to which Chomsky's view of the creative aspect of language use bears on the fundamental issues with which he deals in linguistics, including—perhaps especially —how he thinks of a theory of meaning. Implicit in his attack on Skinner's behaviourism (Chomsky, 1959) but explicitly introduced (in Chomsky, 1966) and variously formulated since, the creative aspect of language use is a set of commonsense observational generalizations that anyone can make, often just by reflection on cases in which one has observed people using language. They are not, then, particularly 'theory-laden'. Using the terminology introduced above, language use is unbounded, stimulus free, but (typically) appropri-

[11] Keep in mind that if my hypothesis that meanings are SEMs or LFs is correct, there are already well-developed syntactic theories in place that individuate meanings. Perhaps a 'physical' (neural?) reduction or accommodation will come some day, for SEM descriptions are event sortals, and SEMs are mental events, but there is no reason to try to individuate by appeal to unknowns something that is already well individuated.

ate.[12] Chomsky insists that linguistics respect these observations: no theory can hope to say what a human language is if it runs afoul of them. Often they play a role in criticism. Take, for example, the observation that language use is unbounded. This implies that a language must provide for an infinity of states. Taken with the fact that these states cannot be dense or continuous, but discrete, it turns out that non-generative theories of language that allow for non-determinate states are out of the question; they cannot provide for the discrete infinity of expressions that provide an unbounded stock to the language user. In another vein, take the observations that language use is stimulus free and yet appropriate. Behaviourism—committed as it is to treating linguistic productions as stimulus bound and to treating 'appropriate to' in terms of circumstance-response mappings to which people have been conditioned—is misguided from the start: it not only abrogates stimulus freedom but claims that there must be some causal explanation for why people say the things they do, given various circumstances. In a related way, we have seen that there is a connection between Chomsky's rejection of an externalist approach to language—represented in the view that a computational theory of language must go outside the head—and insistence on the creative aspect of language use. When you go outside the head by making your theory of language responsible for speaking to the connections between words and things, you make your theory of language into a 'theory of everything'. Finally, when you think in terms of expressions as representations in the sense of re-presenting something in the world, you introduce normative issues: you have to speak to epistemic issues of truth and correctness of judgment and have to contend with the contrariness of human action. All these difficulties, highlighted by the creative aspect, show why semantics as it is often understood—as the domain of reference and truth—is beyond the scope of a serious science of language.

The observations that make up the creative aspect of language use also play a positive role, however, for one can think of the creative aspect as providing a sort of heuristic. On the one hand, you have free linguistic activities of human beings—questions, commands, exclamations, declarations, conjectures, musings, and speculations that are turned to serve any number of purposes. On the other, you have an internal computational engine that provides an endless set of sound-meaning (<PHON, SEM>) pairs, or expressions. How do you think of the relationship between the products of this computational engine and these free enterprises? You know you cannot hope to capture the workings of the computational engine in terms of the uses of language. You also know that you cannot have these free uses without the outputs of the computational engine. So you have to get expressions to 'talk' to systems that allow for the use of language, and talk in the right way. In Chomsky's terminology, you have to think of linguistic expressions

[12] These are the terms on which Mark Baker settles in his unpublished MS 'On the Creative Aspect of Language Use'. Chomsky uses other terms too e.g. 'innovative'.

as consisting of dual 'interfaces' that provide 'instructions'. In the case of the interface PHON, an expression provides instruction to systems for linguistic production and perception (speaking and hearing). In the case of the interface SEM, or meaning, it provides instruction to conceptual-intentional systems. 'Instruction' suggests, as intended, that the interfaces internally constituted in the language faculty configure the other systems, not the other way around. Neither the production-perception nor the conceptual-intentional systems are particularly well understood, although the production-perception systems seem to be better understood than the conceptual-intentional. The cognitive aspect of the faculty of language, or the computational system itself, is by contrast at least partially understood—there is a serious scientific enterprise devoted to its investigation, and with respect to capturing its structure, at least, there has been considerable progress. Furthermore, due to the creative aspect of language use taken as heuristic, we know that this language faculty or computational system must provide a meaning interface or SEM to the user of a language. The creative aspect suggests, then, what a theory of language must provide for in its expressions. It must provide for human 'usable' outputs. In this sense, it speaks from another point of view to the question of the function with which a computational theory of language must deal—as Marr held, something that the computationalist must settle on before choosing formal 'symbols' and algorithms for a domain. It says that the language faculty must provide an output that can be put to human (stimulus free but appropriate) use.[13] It must provide an endless set of perspectives that suit human interests.

3.2 Providing Perspectives in the Minimalist Program

3.2.1 Preliminaries In recent papers directed to general and philosophical audiences (1993b, 1995a), Chomsky has called the meaning interface provided by SEM a 'perspective'. Read as 'a configuring instruction to conceptual-intentional systems', it is a perspicuous term, and should be kept in mind even when tackling the technical discussions of the minimalist program (in Chomsky, 1992b, 1995b). For in order to understand Chomsky's theory of meaning, it is crucial to take into account both his technical work in linguistics and his philosophical writings. I emphasize the philosophical aspects here; my discussion of the technical details is devoted primarily to explaining how the computational system places meanings within words themselves, ignoring other aspects.

But first I need to explain why the term 'SEM' is preferable to 'LF', which has been the term of art for many years. The difficulty is that the gloss for LF, 'logical form', suggests things to philosophers that it should not. It suggests the philosopher's conception of logical form which, as the name indi-

[13] Compare Marr, 1982, p. 31: 'Vision is a process that produces from images . . . a description that is useful to the viewer and not cluttered with irrelevant information.'

cates, is defined in terms of *logical* role—in effect, in terms of role in truth-preserving inference, such as that found in a logic or a theory of the world. Basically, the philosopher's logical form is defined within the domain of intentional content and language use. It is defined in terms of correct use, where the use in question is often assumed to be application of a language to the world. Chomsky's SEM-LF cannot be defined in this use-dependent and externalist way but in terms of intrinsic morphological-lexical features and derived structure. To be sure, this fact about LF is not obvious in received wisdom and in literature, including even Chomsky's. First, many of Chomsky's philosophical interpreters have assumed that LF is also the philosopher's logical form (or perhaps that the philosopher's logical form supervenes on Chomsky's LF); they have tried to assimilate LFs to Davidson's logical forms, and have tried to treat LFs in terms of Davidsonian truth-conditional semantics.[14] But LFs are not defined in whole or part through logic, which is primarily an exercise in correct inference; they are defined over and generated within a syntactic module, and if they figure in language use, it is because they are prerequisites of and provide instructions for use; they constitute the perspectives that can be used in various ways. Second, Chomsky did not state clearly until the advent of the minimalist program that LFs are not only *not* functionally and extrinsically defined in terms of good judgment, good inference, and the like, but that they *are* fully internally defined in terms of intrinsic, broadly syntactic features.[15] Keep in mind, though, that he has for at least 30 years insisted that language use, including reference, is 'creative', and essentially free. And he has for many years insisted on a closely related point: while there can be no proper theory of use, including truth and reference, there is a theory of LFs. These considerations should have scotched the idea that Chomskian LFs and Davidson's logical forms are within the same domain, but it did not and has not done so. And third, there is room for confusion in the fact that LFs are interfaces. Of course LFs are relevant to use; they provide perspectives for and are even suitable for use. They are also the *locus* of 'interpretation', where interpretation is, among other things, a matter of assigning referents, truth values, and the like. Indeed, among the features found at LFs are some that bear directly on issues of logic and inference; for example, LFs specify—syntactically—quantifier scope and provide for something like the variables logicians are fond of. Nevertheless, they are not defined in terms of interpretations; rather, they provide specific and essentially human linguistic

[14] See Lappin, 1991; Higginbotham, 1985; May, 1985; Larson and Ludlow, 1993.

[15] It should be said that *Aspects of the Theory of Syntax* (1965, pp. 79f) places what Chomsky would now call semantic features such as + HUMAN firmly within syntax. But he did not allow—at least in the terminology he used—for the idea that semantic representations ('deep structures', related to 'kernel sentences' (Chomsky, 1957) and T-markers (Chomsky, 1975)) could themselves be meanings. He seems to have thought of meanings as somehow to be dealt with by a 'use theory'—another component to linguistics, but certainly not syntax (1975, p. 18)—and later by 'interpretation'.

structure and texture (*qua* interfaces) to the conceptual-intentional 'cognitive domains' where we use them. If this fact and the fact that they are found only in the head are forgotten, confusion is likely.[16] To avoid confusion, SEM is preferred.

I also need to introduce and explain another important term of art, 'I-language'. An I-language is a language taken to be an object of linguistic science; within the minimalist program, an I-language is fully defined by providing a list of the features that make up an individual's lexicon. To a first approximation, an I- language is an idiolect. More carefully, though, an I-language instantiation in a particular person's head of Universal Grammar (with parameters[17] fixed and a lexicon specified) is individual, internal, and intensional—and one might as well add, intrinsic and (largely) innate. It is *individual* in that it is a linguistic capacity that a single person has. This is easy to see. Assume that linguistic science is a science of expressions. Expressions are sound-meaning pairs. It is very unlikely that any two persons, even within a single language group (English, French), pronounce their words in exactly the same way. Thus, it is very unlikely that there is a pair of individuals who have the same PHONs; if so, their expressions (sound-meaning pairs that they are) differ. In fact, people probably have several different I-languages at the same time; you are likely to pronounce words systematically differently in some contexts than another. Parallel points can

---

[16] Chomsky's recently expressed suggestion that one treat model theory as within syntax can also contribute to the confusion, although once it is pointed out that model theory so conceived does not relate language to the or a world, but only to elements in a syntax-individuated domain, the confusion should disappear. See the discussion of Larson and Ludlow, 1993, in section 4 below.

Incidentally, SEM-LFs do determine truth conditions in a trivial way, and constrain them in another, non-trivial way. They automatically individuate a situation, in a way that I explain below, and they thereby specify what any utterance by a person that is assigned a particular SEM (which is 'interpreted' as having that SEM-description) 'says', so that if that utterance is an assertion and s/he makes reference to some thing, in the world, they individuate how that person describes that thing. If his or her assertive utterance is held to properly describe that thing, it is a correct description.

SEMs non-trivially partially constrain truth conditions in that they provide some lexically and formally specifiable (internally specifiable) 'analytic truths'. These 'analytic truths' are not necessary truths of things in the world, for in fact one can even produce understandable and correct inferences that contravene them. But it is difficult to deny that there are 'analytic truths' determined by SEMs: if I cover a roof with tar paper, I cover its outside, not its internal surface. This point comes up again in the conclusion; it is one way in which internal SEMs 'guide' use.

[17] Think of all natural languages as reflecting different 'switch settings' of a limited number of 'switches' available in Universal Grammar, where a child can with minimal experience set these switches. (Higginbotham suggests the terminology of switches; Chomsky has adopted it in his informal exposition.) If natural languages differ in their syntactic structure in 14 ways, each switch involves two settings, and there is no ordering, there can be 2^{14} or 16,384 structurally distinct natural languages, about 10,000 or so more than it is often thought there are. It is assumed that UG has parameters for phonological features too, but there seems to be little reason to think that there is much semantic variation. See below.

be made about SEM; virtually no one has exactly the same repertoire of SEMs in their operating lexicons as another, hence—again—their I- languages will differ. And one's repertoire changes over time. Notice that the claim here is not that for a given SEM, you can find any number of different PHONs paired with it ('Buch', 'book', 'livre' . . .). This, though obvious enough, is what Chomsky calls Saussurean arbitrariness, which is a different matter. Next, an I- language is *internal* in that it is what one has in one's head, independently of how it is used. Third, an I- language is *intensional* in that it is defined under intension. Intuitively, a (proper) theory of a person's linguistic capacity or language faculty, and only that, will assign relevant features to lexical items and add whatever devices are needed to yield a generative catalogue of the expressions in that language. The theory tells us, then, what that I-language is. Fourth, an I-language is *intrinsic* in that no extrinsic features of expressions are included in characterizing an expression: the ways in which, on an occasion, a person uses an expression are irrelevant to the individuation of that expression. And finally, an I-language is *innate* in that the features over which both PHON and SEM are defined are innately specified. (Remember Saussurean arbitrariness: no *pairing* is innately specified, only the features that define sounds PHON and meanings SEM.) This is not a mere stipulative matter, and it is far from trivial; these features of an I-language have to be innately specified if the theory of the I-language—the object of a proper theory of that capacity—is to speak to the basic explanatory constraint on linguistic science, that of explaining the poverty of stimulus data.

Making I-languages the object of linguistic science represents a radical departure from most assumptions about language and about how to approach language. Most philosophers, for example, have in mind by a language something like a normatively configured set of practices engaged in by a group of people. They think that the right way to get a grip on such a language is to look for the 'rules' that these people follow. This way of looking at a language encompasses functionalist, game-theoretic, behaviourist, and (most) truth-conditional approaches; it should be no surprise that these views of language construe 'grammar' as primarily a matter of correct use, or a semantic (truth-and-reference) matter. Overlapping this concept of a language is language individuated politically: Chinese is what they speak in Hunan province. Both the philosopher's and the politician's languages are congeries or fictions. They are not individuated on the naturalistic principles found in specifying an I-language, but on political, legal, moral, and epistemic grounds. They are the products of convenience, norms, and interest, not discovery: one classifies a person as a speaker-of-L in different ways depending on the interests of those making the classification. These groupings can, of course, appear responsible, at least in intent: Davidsonians and others cluster speakers together by treating them as truth-seekers. But while this may be a responsible normalization, its product (the language L) is nevertheless a fiction, as are its accompaniments—'conventions', 'agreement in belief', 'common behaviour of mankind', 'shared knowledge of the world',

and the like. An I-language is a different animal entirely. It is an instantiation of what Chomsky calls a *natural language*, an instance of Universal Grammar with parametric switches set and lexicon fixed. Crucially, the concept of a natural language is not that of a public language as many philosophers understand it. In defining it there are no appeals to Wittgensteinian 'criteria' for correct use, no communal norms or appeals to mastery in some techniques for using a language. Finally, I-languages, as instantiations of natural language, are not *scientific symbol systems*. Unlike natural language, these—the symbol systems of chemistry, physics, linguistics—reach beyond innate endowment; learning a scientific language is notoriously late and laborious, and people differ greatly in their abilities. Scientific systems are interesting, for unlike natural languages, they actually come quite close to allowing one to define the meanings of expressions in terms of theory-internal truth-role or function. The meaning of a word in a theory is its role in the theory, we are often told, and that is not far off the mark. (This idea is, in fact, built into the assumption that an I-language is the object of the science of linguistics, for an I-language is a language under theory-determined intension, assuming that the theory is descriptively and explanatorily adequate.) Notice that if I-languages are natural languages and scientific symbol systems are not, the usual assumption that scientific languages are continuous with the languages humans naturally speak (and do not learn, but develop), is wrong. There are interesting consequences of this. For one, 'H_2O' is not in the domain of a natural language theory of meaning; to think it is is just confusion. The meaning of the natural language term 'water' is found in its relevant intrinsic features.

If the fact that I-languages are naturalistically (biologically) individuated suggests that philosophers and politicians who speak of language deal with fictions (too often with regrettable consequences), the fact that they are internally individuated suggests useful parallels with other mental sciences. Meanings or SEMs are found within the head and are the subject matter of a computational science of the relevant faculty, a science that individuates and provides for generation of all expressions in an I-language. Similarly, colour—shape configurations are found in the head and are the subject matter of a computational science of vision that individuates and generates all 'visual scenes' available to a person. I appeal to this parallel later.

A third and final preliminary matter: I must say what I have in mind by 'syntax'. Because 'syntax' is a technical term, saying what it means is mildly stipulative, at least initially, but a proposal that hopes for acceptance should speak to the practices, and perhaps the beliefs, of those who call themselves syntacticians. These practices and some beliefs pose the problem for what I propose, which is to include certain semantic features within syntax. Many syntacticians think of themselves as dealing with the purely formal features of expressions, such as V, N, D, VP, and IP, and they typically take their enterprise to exclude semantic matters. ('Formal' here, unlike the 'formal' introduced in the discussion of computational theories, *does* suggest 'structural'.) I suggest that while these syntacticians are right that syntax

does not deal with the kinds of semantic issues that they have in mind (reference and truth), it does deal with meanings. Generally, I suggest that the domain of syntax includes all locally determined, intrinsic features of *linguistic* mental events. If so, syntactic sciences include any disciplines that describe and explain the elements of this domain. As we have seen, the locality condition excludes reference and truth. Locality and intrinsicality together allow for pure formal (structural) features like N; but they also allow for phonological, phonetic, and morphological features. Further, I argue, they allow one to place semantic features in syntax. To soothe disciplinary worries, perhaps it is best to say that all of these disciplines and formal syntax together constitute the field of *broad* syntax; adopting this policy, formal syntax becomes narrow syntax. In effect, the proposal makes broad syntax cover all the domains that are included in a plausible computational theory of an I-language, which seems right.

Many philosophers and a few hard-headed linguists might object to broad syntax on the grounds that syntax does not deal with mental events. But of course it does. It is not just that the trend in linguistics for the past forty years has been to increasingly detach the concept of a syntactic object from the naive view that it is a mere mark on a piece of paper or a mere sound issuing from a mouth, where the 'mere' is supposed to indicate both that syntactic objects are material objects or events outside our heads, and that these objects have no more properties than, say, the buzz of a bee or the scratch on a painted surface. Whatever else a syntactic object is, it is a highly structured entity; in the case of a sentence, it has case assignments, NPs and VPs, and so on, each with their Ns, Vs, As, and Ps. Supporting this disciplinary fact is that there is no way whatsoever to make sense of the idea that some admittedly external distal prompt for an expression—the compressions and decompressions in the air brought about by activity in a larynx, say—*could* have the structure and texture that is characteristic of a syntactic object. If not outside the head, then, syntactic entities that are described by features such as NP must be inside the head.[18] To the response that the sounds we hear when we hear words 'coming from someone's mouth' surely cannot have that structure and texture, there is an obvious counter. They do because the word-sounds we hear are themselves mental entities; they are expressions with highly structured and textured PHONs.[19]

If all this puts phonology, phonetics, morphology, and formal syntax in

---

[18] The argument here is basically Berkeley's, and it is unrefuted. Unlike Berkeley, however, I am perfectly happy to say that there 'are' things 'out there' as described and explained within various physical sciences. And I am also perfectly happy to say that the various branches of syntax are physical sciences, even if they are sciences of what is in the head, for all that 'physical' means is that one has an honest science. (It is difficult to say what else it could mean.)

[19] What about the offended complaint: 'But we hear words issuing from another person's mouth, and surely other people's mouths are outside the head!' The same, however, is true of colours, and they too are inside the head. We need a projectivist account here. See McGilvray, 1994.

the domain of broad syntax, all well and good, but does the theory of meaning, or of SEMs, really belong? Clearly, SEMs will have at least some syntactic features, for they will have the structure of I[nflectional] P[hrase]s (effectively, sentences), and thus NPs, VPs, PPs, Vs, Ps, and so on. The recalcitrant syntactician can admit this without committing him-or herself to including the theory of meaning in syntax: the structure of meanings are included in syntax, it could be said, but not their texture, or that which makes SEMs meaning-distinct from one another. That is, there are any number of SEMs with the structure [S[NP . . . [VP . . .]]] that differ radically in terms of their meanings, so while your proposal includes the structure of SEMs in syntax, it misses meanings. I have to argue that it is possible to distinguish one meaning from another within a mental, intrinsic, locally determined domain.

It helps in advancing this brief to observe that many semantic issues—not reference and truth, but semantic issues nevertheless—have always been included within formal syntax. As Chomsky says, 'Virtually all work in syntax in the narrower sense has been intimately related to questions of semantic interpretation (and obviously, phonetic interpretation), and motivated by such questions. The fact has often been misunderstood because many researchers have chosen to call this work "syntax", reserving the term "semantics" for relations of expressions to something extra-linguistic' (1997, p. 10). Obvious examples include scope and binding matters, c-command, argument structures of verbs, topic and focus, etc. These alone are not enough to distinguish each SEM that differs in meaning from any other, but they suggest the way to proceed. We need to take a further step, which is provided for by the lexicon in the minimalist program. This program includes fine-grained lexical distinguishing features, called 'semantic' features, in the lexicon and the specification of SEMs, which has the effect of distinguishing among SEMs to whatever degree of grain the best theory of meaning insists upon. Including these features and placing them in a generative theory allows the theory to distinguish—exclusive of reference and truth, of course—each SEM in an I-language from any other. It describes what a language can mean, or possible meanings; incorporated in universal grammar, it says what all existing and possible natural languages can mean. This step shifts much of the burden of constructing a theory of meaning to the lexicon. No one has ever doubted that the task of explaining and detailing the *structure* of SEMs is within the domain of narrow syntax, nor denied that there are computational processes that explain how this structure comes about. So there is no doubt that narrow syntax speaks to the *composition* of SEMs. But we also require a field of study that focuses on the fine-grained textural features found in the lexicon—that is, one which details what these features are and how they enter and play a role in the computational process to finally end up at SEM. Call this field of study *lexicology-morphology*. It must be welcomed into the domain of broad syntax if meaning is to be seen as syntactic.

There has been—and for those who have not encountered the idea, may

be—some resistance to the idea that lexicomorphology should be included in broad syntax. Philosophers who worry about this issue at all are very unlikely to think that lexical items have fully distinguishing *syntactic* features, and are more likely than not to think that if they do have anything like semantic 'features', they are assigned externally, or gotten through community training, etc.[20] One reason for this is no doubt that philosophers tend to think of a lexicon as what one finds in a dictionary that reports usage, such as the *Oxford* or *Webster's*, and of the methodology of the lexicologist as that employed by someone who produces a dictionary like this. This is misguided. If a lexicon is to play a role in a computational theory of 'language processing' (and it must, for that is how linguistic theory treats a lexicon), lexical items must have features that are automatically recognizable by the relevant computational system in the head. Perhaps one can say that a dictionary report of community usage can provide hints to someone who has a lexicon, prompting that person to mobilize some cluster of features in his or her head; you have to allow something like this to make sense of how these dictionaries prove useful. But dictionaries so construed are not contributions to linguistics. They come nowhere near stating what those features are, what sorts there are, how they are structured and relate to one another, and *how they enter into language processing in the head*. Linguists too, though, might be reluctant to admit lexicomorphology as a science. While they recognize that a lexicon *must* play a role in language processing, they are not always in agreement about what language processing is, so can disagree about what counts as features of a lexicon. They may also disagree about methodology. And many have the widespread conviction among formal syntacticians that semantics is a matter of relating language to world, not something that the lexicologist can accomplish by detailing intrinsic features within a broad syntax. Given these factors, and the fact that lexicomorphology is very young as a science, it is perhaps no surprise that a fully internalist account of meaning requiring fine-grained intrinsic syntactic features would not strike even many linguists as promising. Surely, however, the right attitude here is this: if Chomsky's minimalist approach provides a programme that is making progress—the empirical proposals that respect the principles of the program actually do meet the major explanatory constraint on linguistic theory (dealing with acquisition)—why not take seriously, at least provisionally, the idea that the minimalist account of lexical items and SEMs—one that introduces fine-grained intrinsic, mental, and locally determined features—is exactly what is needed for a theory of meaning? You not only

[20] Bilgrami, 1992, for instance, seems to think that fine features of the lexicon are inaccessible to the lexicologist, although for reasons that are likely to convince only a philosopher wedded to the idea that meanings really are tied to uses (albeit construed nonnormatively). See also his comments in Chomsky, 1993b. So far as I can tell, one reason he thinks this is that he gives the lexicologist very limited tools: observation of verbal behaviour. This is, I think, completely misguided—a regrettable inheritance from the behaviourist tradition, fostered by Quine on translation.

place your work in a strong programme, but gain the conceptual benefits that that programme emphasizes—among others, concentrating on what can be done within a linguistic science and ignoring what cannot. It would, of course, be foolish to adopt this strategy if there were strong counter-evidence. But there is not. In fact, there is supporting evidence of a very useful sort. Work done on the structure of the lexicon by computational linguists like Pustejovsky (1995) seems to be moving in just the sort of direction that the minimalist strategy for meaning requires. Admittedly very incomplete in the lexical items it has covered, this computational approach nevertheless assigns features to lexical items if various sorts (Ns, Vs, As) that look to be close to what is needed to individuate the 'perspectives' provided by SEMs, and he rightly treats these features as present in a lexical item from the start. To be sure, he adopts a categorial grammar in (1995) and earlier work, thereby treating syntactic processing as aiming towards a sentence to which a referentially based assignment is made, and the minimalist account of meaning can do without that. But this aspect of his account could be abandoned along with some related ones without losing the parts that are tailor-made for a minimalist who wants to construct a theory of meaning within syntax, broadly conceived.

3.2.2 How to Make a Meaning I present here some of the semantically relevant aspects of the minimalist program's picture of the computational system found in the language faculty. Some are beyond the scope of this essay: I will not attempt to explain how technical discussions among linguists bear on meanings in the minimalist program, even though it is clear that they do. For example, in his introduction, Chomsky (1995) ties the interface SEM (and PHON) to biological constraints on these interfaces. This claim is within the scope of this essay, and I suggest a way to understand it below. But I do not try to speak to the way in which the tie he insists on between biology and SEM bears on the difference between what linguists call X^P movement (which includes binding and scope issues) and X^O movement (matters of 'control'). It is enough for my purposes to point to how binding and scope—and, of course, fine-grained features—bear on the requirement that the language faculty be 'legible' to other biologically based cognitive systems.

Second, I cannot provide anything like the detailed cataloguing of lexical features that an approach that requires a completely catalogued lexicon depends upon. I offer a few examples that I think point in the right direction, and focus instead on the basic structural and conceptual issues. There is no lack of detail in the literature, although, as I mentioned, too little agreement on what this detail is supposed to contribute to 'language processing'. The minimalist program's greatest benefit may be that it provides a clear idea of what language processing consists in and how it is best achieved, and thus of how to conceive of the lexicon's contribution to it. But linguists are going to have to come to recognize this.

The minimalist program's picture of language computation is a very thin form of what is called the 'principles and parameters' (P&P) model. The

original P&P model had four levels of representation: deep structure (DS), surface structure (SS), phonetic form (PF), and logical form (LF). Processing within the model proceeds in a 'Y'-shaped way: one begins at the bottom of the Y at DS, goes to the point where branching begins (SS), and proceeds independently to two outputs at the top of the Y (PF and LF). The minimalist program eliminates DS and SS, leaving PF (PHON) and LF (SEM)—that is, leaving only sound and meaning, but maintaining something like the computational model. Minimalism aims not only to increase the importance of PF (PHON) and LF (SEM), but to more closely align these with Chomsky's concept of a biologically specified human nature. The idea, intuitively, is that sounds and meanings are the conceptually necessary parts of any language—what every natural language must have—and we should think of any others that have been introduced at various times, such as DS and SS, as epiphenomena. They, and various apparent rules, principles, and 'theories' that have been stated over them—such as the projection principle, binding theory, and especially case theory—can be seen not as natural objects, but as artefacts of a very simple computational system that succeeds (or does not crash) only if it meets the 'bare output conditions' our biology imposes on the operations of the linguistic computational system. (This is connected to Chomsky's view that if human language systems offer the best way to accommodate these output conditions, they are 'perfect' (1995b, pp. 1, 6–9 221): simple and local in operations while accommodating.) Homologues of a sort for DS and SS remain in the minimalist form of the P&P programme, but DS becomes, in effect, something like the start to operations on lexical items, and SS the place where sound and meaning diverge. Specifically, the computational work other than constituting a 'level of representation' that was done by DS is given to 'Numeration' (production of a list of lexical items with indices that indicate the number of times a particular lexical item appears in the numeration) plus 'Select' (which selects items from the numeration and reduces their index by one). Select begins the computational process, which proceeds locally over features of the list of lexical items and (ideally) over those features alone. The computational work that used to be done by SS is now given to 'spell out': it is an operation that strips the phonetically relevant features from (derived) lexical items (i.e. (complex) feature sets) and directs them towards further processing leading to PHON; the semantically relevant features continue (in 'covert' processing which involves formal and semantic features alone) towards SEM. Note that after spell out, processing diverges; PHON and SEM do *not* 'speak' to one another, period: the total derivational process that proceeds over features of lexical items provides the only link between them.[21]

[21] This point (variously stated at different stages of the development of Chomsky's Universal Grammar) is still ignored by those philosophers, psychologists, and others, who think that the aim of Chomsky's grammar is to provide a transition from sounds to meanings and vice-versa: a sound comes in the ear, and computation leads to a meaning, or one chooses a meaning, and the computation goes in the other direction to yield

The very simple computational process that goes from Select to PHON and SEM consists in two operations that can apply anywhere, at least everywhere on the route from lexical items to SEM. (PHON after spell out requires its own procedures.) One is Merge, which takes pairs (say) of lexical items with their features and puts them together, producing a new syntactic object consisting of the set of features of the pair of lexical items.[22] The other is Move or—preferably—Attract. Attract/Move 'displaces' features to different positions; it leads to treating a noun at SEM as being in the scope of an operator, for example. Attract/Move is the primary focus of attention for the syntactician concerned with what was once dealt with under the topic of movement rules, including transformations (now both long gone).[23] In effect, Merge and Move/Attract operating locally over lexical feature sets provide all the structure, including 'governance' and the like, represented in a phrase-structure diagram and captured by a phrase-structure grammar plus transformations or—within a government-and-binding or an earlier P& P grammar—by various 'theories' and principles, such as Quantifier Raising. Governance, chains, Quantifier Raising and other rules and principles that grammars used to rely upon become epiphenomena of lexical features, the basic operations, and bare output conditions, their roles now taken over by Merge and Move/Attract operating on sets of features, 'trying' to suit bare output conditions at PHON and SEM. In this respect, minimalism advances much further the explanatory aim of linguistic theory: the child need 'learn' almost no 'rules' to acquire an I-language, just acquire some lexical items. For technical details, the best current source in print is undoubtedly chapter 4 of Chomsky (1995b). For very useful discussion of the concept of level of representation and of the transition to minimalism from earlier work within the principles and parameters framework, see Hornstein (1995).

Language processing in this model is clearly modular. It proceeds over language-specific material (features of lexical items) alone in ignorance of anything that may be going on outside the faculty in other cognitive systems, and in the world. All that is required is that the computational process 'converge' in a way that is definable, first, over general economy conditions on

---

a sound. This is a misunderstanding of Chomsky's computational systems, at least. His systems are not 'representational' (starting with a sound or meaning, and moving to the other, or perhaps taking representations of both and seeing if they match), but 'derivational' (1995b, p. 223).

[22] Note too an asymmetry in the processing after spell out leading to PHON and SEM: that which continues to SEM is taken to be like that which occurred before, it meets a 'uniformity condition'. That which leads to PHON could well diverge—presumably, to meet special conditions that the human speech perception and production systems impose.

[23] A central advantage to the minimalist project is that it affords a way to eliminate operations like 'quantifier raising' in favour of simple, entirely local operations on sets of syntactic objects (lexical items and their merged joins). I say no more about this here: it falls in the category of semantically relevant aspects of the program that are too technical to discuss

the computations themselves (effectively, the computation must meet general conditions of simplicity, locality, non-redundancy, and the like) and second, over conditions imposed from without on this specific (linguistic) computational system, where these conditions are localized in the interfaces PHON and SEM. These bare output conditions insist that what appears at PHON and SEM be items that can be used by other cognitive systems. That is, the outputs or interfaces of the language faculty must be *legible* to these other systems, which suggests why 'Attract' is preferable to 'Move': if there has to be 'displacement' or movement at all, it had better be motivated by the need to be legible, or 'visible' at PHON and SEM, rather than by something found within lexical items themselves, which is what Move alone may suggest. Note that there are no other levels, SS or DS, that impose further conditions; any language-specific processing conditions are imposed at interfaces alone. And, ideally, any 'displacement' must be driven by the need to meet the conditions imposed by outside systems at the relevant interfaces—to make language producible/hearable and interpretable by the other relevant systems in the biological organisms we are.

Something Chomsky (1995b) calls an 'inclusiveness condition' provides useful insight into this picture of language processing. Informally, inclusiveness amounts to insisting that everything that comes out of the derivation of a <PHON, SEM> pair—that is, a sentential expression—must be included in the features of the numerated lexical items that went into it. Ideally, no features not present in a numeration are added; they just get merged and moved and in some cases erased. But while one can hope that this condition is met in a Numeration/Select → SEM computation, it is not in the Numeration/Select → PHON computation: phonetic features are introduced. If specific phonological/phonetic computation is isolated after spell out, however, and Merge and Move deal fully and adequately with N/S → SEM computation, we can think of the language faculty as primarily devoted to the production of meanings; sound is peripheral. And if inclusiveness is met, the semantically relevant features of lexical items—those that can play a role in interpretation at SEM—can be thought to survive unscathed and perhaps even untouched until SEM, where they become available for interpretation or, generally, 'use'. Some—those involved in binding and scope at SEM, for example—may undergo modification by Merger or Movement. But if, to a reasonable approximation, semantic features get 'carried along' in a computation and nothing is eliminated that is irrelevant to interpretation at SEM, one can locate a major part of the task of semantics in the job of constructing an adequate lexicon.

What, then, is the lexicon? A lexicon is the collection of lexical items that an individual has at a time. A lexical item consists of three sets of features—phonological, formal, and semantic. In line with a policy he adopted years ago, Chomsky suggests that a theory of the lexicon include only 'exceptional' features—those that are not already specified in UG (1995b, p. 235). He does this to locate all that is special to an I-language, including Saussurean arbitrariness (see below) in the lexicon. However, it is convenient for present

purposes to conceive of appropriately parameterized UG-specified features as included in the lexicon too, so that a parametrically set I-language is fully represented in the lexicon of an I-language. The phonological features are irrelevant in a discussion of meaning; they do not appear at SEM. The formal features are the familiar ones—substantive categorial features such as N, V, A, P—plus functional features, such as T[ense], D[eterminer], and C[omplementizer] that play such an important role in the 'displacement' needed to meet bare output conditions. The semantic features can be thought of—I suggested earlier—as fine-grained versions of the substantive categorial features.

A lexical item must be distinguished from a 'word'. Take, for instance, 'bank'. This is as ordinarily understood an *assembled* item. One way to assemble the word 'bank' is to focus on phonological features of several lexical items alone, ignoring both formal and semantic features; this word contains several lexical nouns and verbs. Or one could take into account in addition lexical and formal features but still not semantic and assemble two 'bank'-words—a V-'bank' and an N-'bank'. While still assemblages, these get closer to lexical items. To get a fully individuated lexical item, go to the semantic features. These semantic features along with the others are sufficient to fully individuate a lexical item—ignoring, of course, any use properties that might arise when a lexical item's formal and semantic features are placed in a SEM and used by a person. The formal vocabulary for these features is not fully developed, although there are some serious and useful proposals, which I outline below. Speaking very informally, I have an N-'$\underline{\text{bank}}_1$' for berms and other earthy projections, an N-'$\underline{\text{bank}}_2$' for financial institutions, an N-'$\underline{\text{bank}}_3$' for the sides of rivers, a V-'$\underline{\text{bank}}_4$' for transcations with financial institutions, a V-'$\underline{\text{bank}}_5$' for berming, a V-'$\underline{\text{bank}}_6$' for controlling fires in a certain way, a V-'$\underline{\text{bank}}_7$' for tilting, as with flying planes, and a V-'$\underline{\text{bank}}_8$' for billiards. Perhaps there should also be an N- '$\underline{\text{bank}}_9$' for fog banks; but no final decision is needed now, for whether this is separate from the earthy projections N-'$\underline{\text{bank}}_1$' can be decided by the relevant best theory of this domain—the theory that adequately individuates lexical items.

Lexical items in an individual's lexicon will always display what Chomsky calls 'Saussurean arbitrariness'. In order to appreciate what this is, keep in mind that the features that define a lexical item are *not* 'arbitrary'. They are drawn from an innately fixed stock of possible sounds (phonological features), and meanings (semantic and categorial formal features such as N). No one I-language contains all the possible and actual human phonological features; an English-speaker draws upon one cluster of the available sounds, and a Swahili-speaker another. And there is variability within formal features too, all provided for within the 'switch settings' (parameters) of UG. Parameters are apparently not needed with semantic features, though; these seem to be relatively uniform in all human natural languages. Perhaps this is related to the fact—Chomsky (1997) speculates—that the child has so much more evidence available to make swift choices in the phonological-phonetic (and formal) domain, so can afford to devote some effort to setting

those switches. Whatever the explanation, degrees of freedom due to parameterization among phonological and formal features is *not* arbitrariness. Arbitrariness consists in the fact that a person's I-language can assign any PHON that is 'available' in that person's phonologically specified cluster to a given meaning or SEM. These pairings are represented in a lexicon. Generally, of course, because a person wants to communicate with others in his or her speech community, s/he will choose to pair sounds and meanings that resemble the sounds and meanings paired in the lexicons of others in community. I am probably well advised to pair to my sound 'arthritis' something like disease of the joints, rather than disease of the large bones (or striding under an oak tree, or anything else). But I need not. There is no natural connection between sounds and meanings, only arbitrary pairings in a particular person's lexicon that feed into a computational process that must meet bare output conditions.

The lexical item is not anything over and above the features that are its constituents. It is not a 'something' in which phonological and formal features happen to 'inhere'. It is just the set of features themselves (and the set is nothing over and above the features, either). Nor is it an abstract object. It is, or is on the way to becoming, a well-defined theoretical entity. In a related vein, notice that assembled entities as introduced above (the single word 'bank', for example) require grouping individual lexical items, perhaps in various people's I-languages, and treating them as the same. If done by appeal to naturalistically defined lexical features (phonological, formal, and semantic) held in common, and that alone, this can be a harmless, and even useful, exercise in generalization: speaking of a population of speakers in which all have set their parameters the same way presupposes it. While assemblages stay within the confines of Universal Grammar by speaking only of features and are plausible abstractions, congeries defined in terms of the uses to which an expression is put are what I called earlier 'fictions'. They are individuated by 'use values'—by the role that a linguistic expression plays in some sort of human enterprise in which it figures. 'Words' put together in this way—however their creation is motivated, whether politically or epistemically—cannot be allowed to play a role in the computational system, nor in the theory (broad syntax) that deals with it. So far as broad syntax and the theory of meaning is concerned, use values are not features of lexical items.

Taking everything so far into account, the thesis is that a meaning is defined in terms of the feature specifications that individuate SEMs within the computational theory of the language faculty as understood within the minimalist program. Thus, a particular meaning is defined in terms of the formal and semantic features that remain after phonological properties are stripped off at spell out and after any other features are eliminated. As I suggested, it is possible to introduce additional properties of *expressions* (<PHON, SEM>s, not SEMs alone) that represent the roles that an expression or class of expressions plays in some enterprise an individual or class of individuals engages in, but these properties of 'features' do not

individuate a meaning. First, they do not explain a meaning in that they in no way explain why an expression has the meaning that it does; only the computational system for an I-language system—operating over a particular set of lexical items—does that, and does so within a naturalistic science of sounds and meanings. Second, they do not capture the essential features of that particular meaning. Only the set of formal and semantic features does that, and these formal and semantic features are drawn—unlike use 'features'—from an innately fixed stock.

3.2.3 The Science of the Lexicon If we are to construct a science of the lexicon, we need a formal vocabulary to refer to and distinguish features. An interesting proposal for the vocabulary of a science of the lexicon—proposed, perhaps, by a discomfited narrow syntactician who is suspicious of admitting special additional vocabulary to deal with the fine-grained syntactic individuation required for a lexical item—suggests that formal feature vocabulary alone could be made to do the work of individuation by appeal to something like the idea of 'frames'.[24] Assume, reasonably enough, that the domain of conventional syntax includes verb 'frames' (e.g. [V NP PP], which in this case is satisfied by any expression that has a verb followed by a noun phrase and a prepositional phrase, as with 'gave a widget to Willy' and 'donated her assets to the Red Cross'). So conceived, formal features can individuate more finely than might otherwise be thought by considering *sets* of frames: investigating further frames in which appear two or more verbs not distinguished by a single frame can reveal that one or more fit some frames but not others, thereby further distinguishing the verbs in terms of conventional syntactic categories. *If* this technique yields a single set of frames for a single lexical item, this set could be used instead of rampant but suspicious terminology like +ANIMATE.

If this proposal succeeded, it would clearly show that all individuative features are syntactic, and intrinsic. But to be successful, it has to show that there is at most one set of frames for each lexical item, and the *prima facie* evidence is that one runs out of at least a conventional catalogue of frames well before each distinct verb can be assigned a unique set of frames.[25] No

[24] The suggestion came from Mark Baker. (He is not a discomfited narrow syntactician.)

[25] See here Pinker, 1994a. Pinker uses this fact to argue in favour of appealing to what he calls semantic information in language-learning, where 'semantic' here amounts to extra-linguistic. The Pinker-Gleitman controversy on the role of syntax in lexical acquisition is compactly and well represented in a collection of articles in *Lingua* that were originally papers given at a conference on the topic (Gleitman 1994; Pinker, 1994a; Grimshaw 1994). Pinker's argument turns on assuming that syntax is essentially abstract and categorial, so that semantics (which is not defined) is required for acquisition of fine-grained distinctions. Gleitman's turns on making syntax less abstract and general, but still sufficiently abstract and general that it does not reach down into nouns (where, she agrees with Pinker, one needs to appeal to semantics). The dispute strikes me as misdirected, driven by a view of syntax that amounts only to a decision to identify syntax with what the narrow syntactician deals with (formal features alone), aided by much-less-than-clear views of what semantics is.

obvious conventional frames distinguish between rinse and wash, for example. More important, however, is that verbs are far better off in this regard than Ns: nouns are not frame-makers. They do not have argument structure but are found in places in argument structure (at SEM), where they with their verb-derived semantic roles (theta roles) are 'interpreted' by using them to refer. So we had better look for another way to construct a vocabulary for the lexicon.

The default alternative to the discomfited syntactician's proposal consists of vocabulary like ±ANIMATE, ±CONCRETE, STATE, PROCESS, and the like, introduced to capture semantic features in the lexicon. This vocabulary is often found in the literature, and it does provide some insight; for simple cases, it may suffice. But clearly, we need great advances in it. One way to advance is to look for organization and structure within this vocabulary, so that the vocabulary and the features it individuates come to be more organized than mere lists of terms might suggest. Some of Pustejovsky's (1995) suggestions are very useful here. He provides for a lexical item a kind of matrix of information that that lexical item contains. Some of the slots in the matrix are hardly surprising. In the case of a verb like 'sink', for example, one slot specifies what Pustejovsky calls its *event structure*. Specifically, 'sink' involves a process, a state, and a specification of precedence: the process precedes the state. Furthermore, there must be a way to assign to 'sink' an argument structure: this verb requires two objects (the sinker (a plane, for example) and the sunk thing (a ship, for example) and will allow an optional third entity specifying an instrument (a bomb, for example). These requirements and options are standard; anyone working in this field knows that this 'information' is contained in 'sink', although there is dispute about how to represent or capture it. The truly interesting contribution of Pustejovsky's view is found in another entry in the matrix, an entry that is itself a matrix derived in part (following a suggestion of Moravcsik, 1975) from Aristotle's catalogue of modes of explanation for things. Specifically, Pustejovsky treats each lexical item as providing information in one or more, sometimes all, categories within the following modes: *formal, teleological* (purpose; what something is for), *constitutive* (what something is made up of), and *causative-agentive* (how something comes into being or is caused). This seems to work quite nicely with nouns; nominal meanings usually speak to these issues. But there are, I think, problems with using it elsewhere. Aristotle took his modes of explanation to deal with things or substances, so it is not surprising that the explanatory mode matrix does not prove particularly illuminating with verbs. In the case of 'sink', Pustejovsky suggests (1995, p. 192), the *agentive* specification is sink qua act, and the *formal* sink qua result. It is not clear to me, at least, how 'agentive' and 'formal' attached to two different constituent event parts of sink advances the task of individuating sink beyond what is already provided for by a specification of event structure (which is not much of an advance: this particular form of event structure does not distinguish sink from thousands of other verbs, including rinse and wash). With nouns, though, this explanatory mode matrix is very useful indeed. The entry for the noun item 'book', for

example, includes these feature specifications: (1) an *agentive* specification in write, (2) a *telic* specification in read, and (3) a *constitutive* specification in information · physical object. The last is a 'dot product' type of classification; intuitively, a book, like other printed matter such as a newspaper or dictionary is *both* physical object and information. It is also something that is written by people, and that people read. Undoubtedly more needs to be said here, but this seems to me to be progress: we can begin to see how to individuate certain lexical items in the same formal category. We can also see how, if this information is a part of the meaning of the word, when a person uses it, s/he can focus on one thing rather than another, and sometimes both: 'The book that Harry is writing [information] will weigh a kilo [physical object]'. What, then, of (4) Aristotle's 'formal' specification? With nouns Pustejovsky says that this specifies how the lexical item fits into (is bound by) the 'argument structure' of the lexical item. Specifically, the formal specification for 'book' says of a book *qua* physical object that it holds or contains information. Intuitively, and without the machinery of argument structure (which looks strained when asked to do duty for nouns not derived from verbs), this is a way to indicate that books, like newspapers and magazines, are within the genus printed matter. This, presumably, is close to what Aristotle had in mind by a *formal* mode of explanation. A better account might, however, be that the formal specification bears on how this lexical item contributes to the computational system.

The details of what appears to be a developing science of the lexicon are, of course, crucial; but this is not the occasion to pursue them. It is not even the occasion to explore how to adapt Pustejovsky's observations and some of his theoretical framework to the entirely syntactic enterprise found in the minimalist program, although the discussion of SEMs or meanings as perspectives below hint at what needs to be done. Instead, I will speak to what I suspect is by now a growing doubt about what I call a science of the lexicon. I assume that the lexicologist-morphologist who pursues the principles of the minimalist program aims to construct a theory of the texture or fine-grained detail found in semantic features excluding—of course—any use 'features' that might accrue at SEM. The problem lies in what appears to be a stark contrast. There is a well-developed theoretical vocabulary to deal with what can be called the coarse features of lexical items; these are the formal features, and narrow syntax provides the vocabulary to deal with them. This vocabulary even *looks* theoretical: 'NP' and 'V—NP—PP' have a comforting resemblance to terms like chemistry's 'HOH'. The doubt I have in mind is raised by the fact that—at least so far—creation of theoretical terms to deal with lexical texture is largely—as a glance above indicates—a matter of taking orthographic prompts for natural language expressions (wash, animate, holds, information) and changing their fonts (ANIMATE, **animate**) or, as I have done all along, underscoring them (information). '*This* is theoretical vocabulary?' the dubious will ask. The response is: yes, and perfectly good too.

Underscoring and font-changing are ways to mobilize the referential technique of *exemplification*. Goodman (1968) outlined an account of exemplification. Exemplification consists of a dual referential relationship: something

(here, an expression in a natural language) is used by a person to refer to a 'label' that in turn is used by that person to denote the thing. Familiar cases include the tailor's swatch used to refer to the 'label' 'mustard yellow with blue pinstripes', which in turn is used to denote the swatch. Less familiar cases include theoretical uses; an experimenter in colour science might take an entity that causes a colour experience in her head and use that colour experience to refer to some label in a theoretical symbol system—'<Hue x, Brightness y, Saturation z>'—that she in turn uses to denote the colour experience. The lexicologist can do something like the colour experimenter does, but with an important difference. Clearly, the lexicologist—and the would-be lexicologist—can take the expression that an underscored orthographic prompt creates in the head and have it refer to a theoretical label. This is what the colour scientist does too. But, unlike the colour scientist and also unlike the narrow syntactician, s/he seems to have no well-developed theoretical vocabulary to refer to, and thus no theoretical vocabulary that can in turn be used to denote the expression. Surely, however, s/he can turn the natural language expression itself into a theoretical term, and have it denote the features that constitute its SEM. S/he can use the expression in two ways—to refer to a label, and as a label to denote the expression. Call the dual referential relationship that involves a use of a single expression in both an exemplifying and theoretical denoting way *sophisticated autoexemplification*. This referential strategy, I suggest, is exactly what the minimalist lexicologist utilizes; in fact, it is what anyone with theoretical intentions or hopes who relies on prompts for natural language expressions utilizes—for example, Pustejovsky in providing the prompt information, Fodor (1982) in providing REFRIGERATOR, or—I suspect—anyone reading this paper. I am suggesting, in effect, that exemplification figures in every attempt that has so far been introduced to construct a science of the semantic features of lexical items. If this is right, not just the minimalist lexicologist, but others, had better welcome sophisticated autoexemplification to their theoretician's toolboxes.

Sophisticated autoexemplification (SA) can easily be supplemented with more comforting vocabulary by appealing to the narrow syntactician's theoretical terms and diagrams, providing what I called complementary means to refer to a SEM. Because SA relies on a SEM to refer exemplifyingly and (in turn) denote, and because each lexical item and any SEM includes formal features, it is easy enough to use SA along with standard narrow syntactic theoretical orthography and depiction. For example, take information and add brackets and the prompt 'N': [$_{N}$ information]. Where you have an expression with structure such as under the table, nest brackets: [$_{PP}$[$_{P}$under][$_{DP}$[$_{D}$the[$_{N}$table]]]. It is possible that there will some day be a completely articulated theoretical vocabulary to deal with semantic features, where this vocabulary is not exemplificational in any way. Certainly Chomsky has been looking for such a vocabulary, at least since *Aspects of the Theory of Syntax* (1965). If we are ever in such a position, we can simply denote semantic features, without also exemplifying them. At the moment, however, this is

only a hope, and given the extent to which exemplification appears in technical vocabulary like +ANIMATE and information, it is probably wise to make sure that it is properly understood, and properly used.

3.2.4 SEMs as Perspectives The view that meanings are in the head and syntactically individuated has important cognitive consequences. The best way to see what they are is to look at Chomsky's view of a SEM or LF as a *perspective* (1993b, 1995a), an adaptation for a general audience of the technical syntactically defined concept of an LF or SEM as an interface. He points to a way in which SEMs are both *rich* and *anthropocentric*, and suggests a way to conceive of how they work cognitively in 'interpreting' the world (1995a, p. 20):

> [SEMs] focus attention on selected aspects of the world as it is taken to be by other cognitive systems, and provide intricate and highly specialized perspectives from which to view them, crucially involving human interests and concerns even in the simplest cases.

I have in mind by richness the idea that SEMs provide very fine-grained media for all sorts of enterprises, not only describing and explaining the world (although not as a science that uses vocabulary apart from natural language would), but writing and reading literature, chatting with friends . . . Calling these perspectives anthropocentric points not only to the fact that linguistic interfaces are specific to humans and within the human mind, so different from any interfaces the cognitive systems of other organisms might provide, but to the fact that they can serve human cognitive tasks. One way—the only acceptable way—to view how they can serve cognitive tasks is to construe them as constituting the intrinsic contents that, when used to serve a cognitive task like description, constitute 'how those things are seen' within the common-sense framework. Looked at in this way, a perspective provides the *linguistic* (not visual, e.g.) content ascribed to a thing, and when the ascription is held to be correct, this content constitutes that thing in a way that is efficient and effective for the organisms we are. In effect, this is a constructivist, projectivist way to construe the use of a perspective, or SEM interface: as with Goodman's worldmaking, human beings 'make' their world(s) (cp. Chomsky, 1997). In sharp contrast to Goodman's view of the way in which humans make worlds, though, here the intrinsic contents or internalist perspectives which are the tools and instruments of worldmaking are internally constituted and largely innate. Indeed, it is only because they are so that we make worlds so readily, and share a world as much as we do. We can, and do, because we do not have to also construct our perspectives. A computational engine in the head provides them—rich, suitable for us, and in endless numbers. That is why it is possible to share as much as we do a common-sense world (the symbol systems of science are irrelevant,

remember) that is readily perceived and effectively 'individuated' for our purposes.[26]

3.2.4.1 Perspectives in Cognition It should be clear by now that SEMs, inheriting all the semantic features of lexical items as they do, are replete with textural detail, so I need not press the point that they are rich. Instead, I will expand on their anthropocentric cognitive role. To get a grip on the general cognitive picture being drawn here, it helps to develop briefly an analogy between the language faculty and the perspectives it provides and another faculty in the head, vision, and the perspectives it provides. We have already seen that vision, like language, is a faculty that is nicely captured with a computational theory; we have also seen that vision, like language, provides rich interfaces that can be described and explained with the appropriate computational theory. In the case of vision, a particular interface can be thought of as something like a visual field at a time; it is described by assigning to each smallest (according to the theory) position in a scene a sextuple of hue, brightness, saturation, altitude, azimuth, and depth numerical values; the computational theory explains how these come to be assigned, given arrayed input. In the case of language, there are two interfaces, but only one, SEM, is of concern here. Any particular SEM is adequately described by a formal and semantic feature assignment, and the computational theory explains how these assignments come about from sets of lexical items. So far in developing this analogy there has been no effort to move outside each particular faculty. To do so, we have seen, is to move outside of the domain of the science of that faculty. But, clearly, humans do 'use' colours to characterize things in a world, and they obviously can use SEMs to describe things. Focusing on colours, which are among vision's outputs, we know that there is good reason to think that they are in the head. But they seem to be 'on' things outside of us. So, the question: how can colours in the head appear on things 'out there'? The answer: by projecting them. As Goodman pointed out, the root of the concept of projection is attribution. We project when we (correctly, by the standards of the relevant discipline or framework) attribute. This general characterization suits both vision and language, although there are obvious differences between attribution in vision—a more or less devoted process that often serves a single task (gathering information about an environment)[27] and that is almost beyond

[26] Chomsky, 1997 remarks that the identities of 'things' (the things of the common-sense world, and one might as well add events too) are fictional. This was a common view in the seventeenth through eighteenth centuries. Hobbes and Locke are clear cases. In recent times, Sellars held this of the things of the 'manifest image', also identified by him with the non-scientific but practically necessary common-sense framework.

[27] On the best current hypothesis, colours provide information about surface spectral reflectance properties of entities as described in an extension of wave mechanics. Nothing like this kind of claim is possible for language, of course: the claim that language provides information about the world is virtually empty unless 'world' is specified. When it is given its usual 'definition'—the (common-sense) world of people, artefacts, water and such—the projectivist points out that this world is far from being the subject

our criticism and control—and projection in language, which can be used for all sorts of cognitive tasks, where epistemic criticism is relatively easy, and where humans can maintain control over use. But the same cognitive picture depicts either form of projection: a rich content inside that constitutes a perspective comes to seem to be 'outside' when it is taken to 'interpret' correctly.

It is also useful to help oneself to a view found in traditional conceptualism, that the forms and properties that we attribute to things in the world are actually not located 'in' things outside, but in the head, where when applied to things they *configure* the things to which they are applied. Similarly, one can think of SEMs as containing the properties of things in the world. They effectively say what persons and water are, providing something like identity conditions on entities when applied to (attributed to) them. We have SEMs in our heads that, when correctly (by our lights) used to describe the things of our world, effectively 'make' those things suit the features of SEMs. These entities are, strictly, artefacts of our cognitive faculties put to the use of dealing with the world, but very useful ones indeed. They are the entities of the common-sense world. Notice that there are truths and falsehoods concerning them, although these entities 'made' by relying on our anthropocentric concepts are not the entities of science.

In fact, because SEMs perform the tasks that traditional conceptualism wanted concepts to perform, they can and should be treated as sophisticated replacements for concepts. Thinking of them as replacements for concepts depends on accepting a deflationary translation scheme that takes various kinds of concepts as traditionally understood (that is, as mental entities separate from language) and recommends treating them as nothing but instances of certain structurally similar classes of the features that constitute SEMs; in effect, the translation scheme proposes a kind of reductive identification of concepts with SEMs. This is not theoretical reduction of that rare and celebrated kind secured now in parts of chemistry and physics. We are not dealing with two sciences, but only one: while 'concept' is a technical term, it can hardly be said that it is well defined in an honest scientific theory. Features, feature sets, and SEMs, however, *are* well-defined 'entities' in the honest science of broad syntax. The proposal suggests that because these scientific entities when appropriately used configure in the way con-

---

matter of physics; it is, rather, a creation of our own. Chomsky would then point out that it is foolish to think that the language faculty is 'devoted' to providing information even about this world.

Incidentally, there are 'colour objectivists' (or perhaps realists) who celebrate the hypothesis that colour vision is more or less devoted to gathering information about surface spectral reflectances (SSRs) by calling these SSRs colours. This is, I think, an extraordinary proposal. Quite apart from the fact that colours as experienced (captured in a science of hue, brightness, saturation triples) are ordered in ways completely different from SSRs, it suggests that the correct science of colours is a rather complicated extension of wave mechanics. Some rather naive physicists might welcome this; practising colour scientists would be astounded. See McGilvray, 1994, and Clark, 1993.

cepts traditionally understood are supposed to configure, they are reasonable candidates for what philosophers were getting at all along when fumbling about trying to say some systematic things about the mind, what is in it, and how it works. The translation scheme is defined over kinds of concepts. A particular *situation-concept* (some might prefer 'eventuality-concept') such as Harriet slept on the porch last night can be treated as nothing but a complete sentential SEM feature set occurring at a time. Thinking then of sentential SEMs (features that they are) as consisting of various phrasal parts, some of which themselves have phrasal parts, property-concepts and other kinds of concepts can be treated as nothing but feature sets for phrasal parts of full sentential SEMs. Mobilizing the theoretical descriptive device of a phrase-structure diagram, one can say that each VP, AP, or PP node in a phrase-structure diagram is a mental event sortal that, when provided with spatial and temporal coordinates for someone's head and correctly applied to a particular mental event, denotes and describes a particular feature set into which a *property-concept* or *relation-concept* can be deflated. Examples include [$_{VP}$carried the trash bin to the end of the road] and [$_{PP}$under the table in the room]. Similarly, each NP node in a phrase-structure diagram can be provided with spatial and temporal coordinates and when correctly used, etc., denotes and describes a particular feature set into which a *class-concept* (note: *not* a class-*description*) or an *individual intension* can be deflated. Generally, one can think of a particular SEM or a relevant part of a particular SEM as a concept instance. If so, we can treat any *linguistic* concept as a SEM mental event particular.

Once deflation of philosophically provocative entities has begun, it might as well continue. The nominalist will point out that terms like 'property-concept' suggest that properties are the sorts of things to which one refers when using a property-concept (after deflation, a VP, PP, or AP). So, s/he suggests, why not drop the offending '-concept' and declare that a *property* is nothing but a VP, PP, or AP? This suits the nominalist's aims: it makes properties nothing but syntactically defined entities. And, so far as I can see, there is nothing wrong with it. Deflation has to be restricted to the properties found in natural languages, for SEMs are defined only over natural languages, and nothing is said here about property-like entities found in the outputs of other faculties (sounds, shapes, colours), nor about anything that might be thought a property within a scientific symbol system. Furthermore, the thesis should be extended beyond properties to situations and individual intentions. They are nothing but sets of features that make up SEMs and relevant parts of SEMs. Thus, the conceptualist-projectivist view that configuring interfaces are in the head favours nominalist deflation. Consider the parallel colour case. Colours have hue-brightness-saturation ordering, and nothing physical outside the head has that ordering. Situations have formal and semantic features: they have argument structure, they are perfective or imperfective, they have tense structure, they have all sorts of textural detail that nothing in the worlds of physics or chemistry has. There is, of course, a sense in which these features are dealt with within biology, but the only

developed branch of biology that deals with them at the moment is broad syntax construed as above. In any case, SEMs in the minimalist program support both deflationary moves,[28] and I suspect that the conceptualist and nominalist should welcome the program and its results: not only are they vindicated, but they can look for other issues to resolve—or perhaps they can become syntacticians, especially lexicologists-morphologists, and contribute to an internalist theory of meaning.

Summarizing this overview of the cognitive consequences of the minimalist program's view of meanings in the head, or SEMs: SEMs in the head get transported 'outside' when persons use these entities to configure 'experience', but really they are in the head alone. This suggests a methodological point: while a computational science of these internal entities may need to proceed by observing how SEMs are used by people—and by appeal to medical data, psychophysical results, statistical analysis, PET scans, specific cognitive defect data, and any other data that one can find—the target is internal events that configure uses. One may need to look at 'effects' to reconstruct the cause, and where creativity plays an important role, this can be difficult. But there is no reason to think the project impossible. In fact, there is considerable progress.

3.2.4.2 Perspectives as Interfaces Recall from the summary of some of the technical aspects of the minimalist program that SEMs, like PHONs, must meet certain constraints. They must, of course, be readily 'reachable' from a selection of features at/in numeration and select. This is a way of saying that the computational procedure must economically and efficiently produce a SEM, or else crash. They must also meet bare output conditions, the biologically based demands that the outputs of the computational system be 'legible' to other systems—certainly the 'performance' systems, but others too, including other cognitive faculties, such as vision.[29]

Explaining legibility by appeal to the formal story alone quickly leads to very technical issues, so to avoid the technicalities, I exploit the idea that if SEMs are interfaces, they must be a locus for interaction with other cognitive systems. In the case of the PHON interface, the identities of at least some of these is reasonably obvious, even though little still is known about them; they are audition and speech production. In the case of SEM, it is not even clear what the relevant systems are, and it is easy to go wildly wrong by

---

[28] Chomsky 1997, outlines parallels between PHON and SEM to make the point that as there is no reason to treat PHON as a mere placeholder that points to something elsewhere in the head or world that does the real phonetic work (a 'lingua phonetica', 'phonetic values'), so there is no reason to treat SEM as consisting of (say) formal features alone that point to something elsewhere that does the real semantic work (a 'lingua mentis', 'semantic values'). These observations also make a strong case for the feature-based computational model built into the minimalist program.

[29] 'Legibility' and 'bare output conditions' are recent efforts to speak to a continuing theme. For 'legible' one can substitute the earlier 'visible', so far as I can tell. And BOCs are clearly related to the principle of full interpretation in Chomsky 1986.

putting the wrong label on them—e.g. 'central processing unit(s)'. Chomsky labels them in the most general way: they are the conceptual and intentional systems. Thus, SEMs are conceptual-intentional interfaces. One could as well say of these systems that they are 'whatever else is in the head that contributes to speech use, or interpretation'. Call them the 'interpretive' or 'use' systems. Think, then, of SEMs as the loci where meanings bear on interpretation or use. SEMs are 'read' here.

It is not crucial to this conception of legibility that one know what the other systems are, other than that they are involved in interpretation and must operate over whatever SEMs intrinsically provide—a set of features. Knowing this, we can look to what is intuitively involved in interpretation—truth and reference, among other things—and ask what could plausibly be fully explained and described within the computational (local) theory of the language faculty and yet clearly play a central role in truth and reference. Examples include a long list of structural contributions (argument structure, quantificational structure, tense structure . . .) plus all the detailed texture provided by the semantic features found at SEM. In my terminology, the language faculty contributes intrinsic contents (variously structured and textured situations, properties . . .) to the interpretive systems that use them. Meeting bare output conditions as they must, these situations are suited to creatures like us, with the systems that we have. They provide distinctive, linguistically configured perspectives 'designed' for use by creatures like us. Suiting bare output conditions as they do, it would hardly be surprising that they met human interests including—it must be emphasized—the need for free and creative activity. Unlike vision's, these are not devoted perspectives.

This is a rather thin constraint, and deliberately so. Keeping in mind that it is plausible to assume that semantic features come along with any surviving formal features as free riders in the computational elimination procedure that leads to SEM, ask whether bare output conditions say something about what these semantic features must look like to enjoy a costless and probably undisturbed trip to SEM. Not much, I suspect; the constraints placed upon what appears at SEM by bare output conditions certainly concern functional features like TENSE and CASE, but it is not clear that they need concern the fine-grained semantic. Perhaps this is as it should be. You do not want to build into bare output conditions a way to rule out well-computed sentences such as 'Colourless green ideas sleep furiously' which do not have any obvious immediate use or interpretation. You want them to count as structurally acceptable perspectives that someone could use for some purpose, even if you cannot imagine what it would be. You want them to have a meaning, even if no apparent use. Otherwise, you impose biological constraints on language use which undermine the expressive range needed for the creative aspect of language use that is distinctive of humans.

Clearly, though, this thin constraint cannot be all that Chomsky has in mind when he says about SEMs that they are suited to human interests ('crucially involving human interests and concerns even in the simplest cases'). The phrase 'the simplest cases' and the examples he provides suggest

another route that makes sense of how SEMs suit human interests: look to the semantic features of lexical items. To see what is in the lexicon and the clustering of semantic features in the lexicon that speak in another way to human interests or anthropocentricity, recall that a lexical item is made up of phonological, formal, and semantic features. The phonological lead to PHON; the formal and semantic lead to SEM. The PHON features are drawn from an innately fixed set, parameterized to an extent so that the sounds that an English speaker will accept as 'sounding like English' are unlike sounds that a Japanese speaker would accept.[30] The formal and semantic features—the only ones relevant to SEM—are also drawn from an innately fixed set, and with respect to the formal features alone there is plenty of evidence of parameterization, whatever 'switch settings' make for differences between 'head first' languages like English and 'head last' languages like Japanese, and perhaps even for differences in capacity to 'incorporate' morphologically (Baker, 1996), will be found in formal features in an I-language's lexicon. It is difficult to say whether there is parameterization in semantic features; for a clear answer to that, we will have to await more nearly complete theories of the lexicon. So far as I know, however, there is no strong evidence at the moment that there is parameterization of semantic features, so I will assume for present purposes that they are not. With some idea of what is fixed, then, let us look at what is not fixed. One thing not fixed is the pairing of formal-semantic features with phonological—that is, the lexicon displays Saussurean arbitrariness. Another is exactly which lexical items will be in a person's lexicon at a given time. As a consequence, it is not fixed just which semantic feature clusters will be found in a person's lexicon at a time. But selection of lexical items is not completely random: there is considerable overlap in semantic clusters among the 30,000 or so lexical items that any child of 6 or 7 has available, for there is plenty of overlap in the discriminations, using language, that every child makes. This begins to bear on the issue. To a good approximation, every child has not only mommy and dog, but give, wash, through, stay, up, good, hold, and thousands of others, although of course the semantic features exemplified here with English orthography will be lexicalized (given formal and phonological features) differently in other languages. Now, it should be obvious that *these* semantic feature clusters are useful, and serve human interests. This surprises no one, though, for no one denies that 'words' like these are useful. The important issue is how they come by their usefulness. For the externalist, it is because the contents of words (assuming the externalist has an account of words) comes from their use by human beings in the first place. For the nativist internalist, it is a matter of innately fixed semantic feature clusters—parts of perspectives that they can be—being intrinsically suitable for crea-

[30] A useful resource here is work done by Jacques Mehler and associates. See, for example, Bigelgac-Babic, Bertoncini and Mehler, 1993; Christophe, Dupoux, Bertoncini and Mehler, 1994; Mehler and Christophe, 1994.

tures like us, or 'defining *our* perspectives'. These are Cudworth's innate ideas, updated and dealt with in a branch of biological science, lexicology-morphology. Unlike Cudworth's story, these are not provided for our benefit by God, but—we assume—by evolution.

The key to 'crucially involving human interests and concerns' seems, then, to lie in the semantic features of nouns and verbs commonly encountered in the lexicons of children and adults, such as 'water' and 'give'. For the internalist, these detailed features of intrinsic contents are fixed innately, for this is the only way to make sense of the fact that children acquire lexical items as quickly as they do, and to make sense of the fact that what they acquire is uniform in semantic features across the whole human population. The strength of these considerations is increased by simple observations that show what some of these features must be. Consider 'The water from the Little Branch rose by the hour and threatened Northville; it already covered the surrounding hamlets. It did not look as if anything in the village would be left after the flood, and the few remaining residents wondered whether to move all they could salvage and rebuild Northville on the north bank.' As this simple story indicates, water threatens, so water must include +ANIMATE and +AGENTIVE; it also covers hamlets, so is +CONCRETE; if Harry sprays water on his plants, though, it is −ANIMATE. As Chomsky (1995a) points out, water transforms when subjected to some simple operations: putting a tea bag in a cup of water makes tea. On the other hand, something that comes out of the faucet may have the colour of tea, and even the chemical composition, but it is water; humans transform water to suit their purposes. Furthermore, one must specify somewhere in the semantic feature set for water something like CONSUMABLE, LIQUID, and other features: like other nouns, as Pustejovsky's account suggests, a lexical feature specification must speak to tells, material constitution, causal factors, and perhaps Aristotelian definition, or definitions. Syntactically complex expressions suggest even more features to deal with, all nevertheless automatically recognized by the child. These N-PP structured concepts, boat in the water, fish in the water, boat on the water, fish on the water, are obviously distinct in ways that depend essentially on the semantic features not only of boat and fish, but of water, and there is an indefinitely large class of other examples with water that make the same point. As for Northville, the name of a village, it inherits the semantic features of village (not city or town, but no doubt overlapping in features with both). From the example, these must include +CONCRETE (what the water would engulf) and +ABSTRACT. These are *both* within the same semantic features set; if they were not, 'village' (and 'Northville') would have to be ambiguous, and there is no reason to think that it is. Now, given that there is good reason to think that these clusters of semantic features are as complex as they are and uniform as they are across the human population (however water and village are lexicalized), it is inconceivable that a child could acquire these words as quickly as s/he does, virtually automatically, unless these were innately determined.

A glance at feature clusters for words like these and other nouns that

appear in natural language lexicons, not to mention the feature clusters for thousands of verbs like 'give' and 'wash', indicates—it seems to me—that the semantic feature clusters within water, village, and very large numbers of others, are 'suited to human use'. One aspect of this claim is that it is very difficult to imagine our water serving any other organisms but ourselves; Martians don't—perhaps can't—have water (nor, hence, water). Another is that tacit, or implicit innate knowledge of these features is required for humans to develop any other kind of 'knowledge about the world'. This point was made by Herbert of Cherbury (1937) when he wrote in 1622 that without innate knowledge, (explicit, articulated) knowledge of things outside, not to mention good and bad, would be impossible.[31] Taken with the first observation and put in the nativist version of a Goodmanian constructivist framework that I have been encouraging, this suggests that 'the world' that all human beings construct on the basis of innately specified features is *our* (human) world—a world that suits human cognitive needs and capacities. The water in it is *our* water, meeting our 'measure', as Cudworth might have said.[32]

3.3 Summary

Meanings are instances of event sorts that are in the head that constitute interfaces/concepts; they are syntactically defined and individuated. They are SEMs, or relevant parts of SEMs. After nominalist deflation, they are situations, individual intentions, properties, relations, and so on. Instances of each of these kinds of entity are identical with wholes or parts of SEM instances. 'X is the meaning of Y' amounts, then, to 'X is the SEM of the Δ-node <PHON, SEM> Y for I-language Σ'. X is a property, situation, individual intension, relation . . .—some relevant part of a SEM, including a node designatum and all features within it. Y is the <PHON, SEM> pair of which the node SEM is a part. In an obvious way, then, meanings (including properties, situations, and relations) are constituents of expressions. The

---

[31] See note 1.

[32] Externalists hope to undermine this idea by trying to convince us that things 'out there' of which we clearly have no innate knowledge play a crucial role in determining the meanings of words, Putnamian, 1975 twin earth thought experiments play a prominent role in this effort: we are asked about 'what one would say' when put in a world that has a substance (a 'natural kind') that looks like water, but has the molecular structure of XYZ. The 'results' of these 'experiments' are very hard to take seriously, particularly if there is good reason to think that scientific symbol systems ('H_2O') are discontinuous with natural languages ('water'). But for what it is worth, my intuitions tell me that if diamonds were made out of very hard H_2O on twin earth and I were transported there, I would call them diamonds, not water, and I would call the stuff I drank from a glass and used to wash the glass water—and this even if the glass happened to be made out of slightly less hard H_2O. Chomsky (1995a, 1997) engulfs Putnamian intuitions with multiple examples, to good effect.

upshot is that meanings or intrinsic contents are in the head, and only there: they are identical with SEMs and the subject matter of syntactic study. Goodman was right: there is a form of syntactic classification that yields contents. And even his constructivist programme has begun to look interesting, so long as it is divorced from anti-nativist prejudice.

4. Going Outside the Head

Virtually everyone today who deals with meaning or other semantic concepts such as reference, interpretation, or content thinks the real issues of content and meaning arise in speaking to how a syntactic object like a SEM is interpreted—in how it is related to the world. I try here to speak to how interpretation looks in the constructivist picture that I have been developing—where interpretation involves a person using intrinsic contents to serve their ends and, in doing so, 'making' a world—by speaking to how it compares with two other pictures that differ, at least superficially. In one picture, Chomsky's SEM-LF is interpreted by assigning Davidsonian-Fregean semantic values to nodes. I explore this option by looking at a single work in formal semantics that appeals to Chomsky's LF (Larson and Ludlow, 1993). I argue that to the extent that these semantic values are made into entities in the world, you get no *theory* of interpretation; if you want an honest theory of interpretation, you must place its 'semantic values' in syntax, not in the world. In the second picture, characteristic of most philosophical work on meaning and content, SEMs are left out—perhaps through ignorance—and word aggregates (not SEMs) are assigned contents or meanings by specifying the roles they play in how humans use them. This basically Wittgensteinian picture can provide no theories, but it can provide more interesting descriptions when supplemented with SEMs. Generally, introducing SEMs should lead to rethinking both pictures—to abandoning one and redirecting the other.

4.1 Formal Interpretation with LFs: Larson and Ludlow

Larson and Ludlow (1993) hold that a theory of interpretation is a theory that yields, for a language L and a set of expressions <PF, LF>s for that language, a further set of LF-semantic value pairs. For each <PF, LF>, the theory yields an I (for 'interpretation'). A theory of interpretation is thus a theory of <LF, I>s: it is a theory that assigns to these new pairs a full description and explains exactly how each is 'generated' from primitives, in the same way as a theory of <PF, LF> pairs describes and generates fully in its domain. Working within the Davidsonian branch of the Fregean tradition, Larson and Ludlow treat the theory that deals with interpretations as providing an *assignment* of objects in a domain and truth values to the appropriate elements of an LF. The aim, intuitively, is to have NPs assigned individuals

and classes, sentences assigned truth values, and other nodes other appropriate 'values'. Interpretations, then, give LFs semantic 'values', and the core of the theory has to be a matter of describing and explaining how they get (on the Davidsonian approach) things and truth values. For a useful image, think of the result of interpretation as a decorated phrase structure diagram: as a Christmas tree gets decorated by adding things to it, so a phrase-structure diagram gets things in the/a world and truth values attached to appropriate lexical items and nodes. Things and sets of things are hooked to lexical items under NPs, sets and pairs (etc.) of sets of things to lexical items under VPs, and truth values to Ss. The crucial question is what the nature of the hooks is—the nature of the assignment hooks that tie things and truth values to nodes. On the Davidsonian approach, the objects that serve as semantic values are existing entities in the world (to meet extensional constraints) and genuine truth values (not arbitrarily assigned values). The aim must be, then, to produce a serious theory of these pairings between LFs and parts of LFs and things, one that tells us what values are and how they are assigned.

The only plausible candidate for such a theory, surely, would be a computational-causal theory that details the links between LFs and things in the world (and truth values?). If that is what it is to be, the theory must in this head-to-world domain say exactly what values are, and how to produce the pairings. If so, Larson and Ludlow and others engaged on the Davidsonian enterprise will not succeed. As I said before, if a computational theory of interpretation is—unlike Marr's theory of vision or Chomsky's morphological elimination procedure (1995b)—committed to dealing with a domain that goes outside the head to include things in the world plus truth values (whatever they might be), there is no such thing at the moment, and the project appears unworkable. As a computational theory, it must still operate formally in the sense specified earlier, and must almost certainly rely on local operations alone. There is no reason to think that there is or could be such a theory. For practical reasons alone, no such theory appears to be in the offing—especially once one takes into account the fact that languages allow one to refer to a global domain. And when one adds the creative aspect of language use, the project seems impossible. Someone might try to *tell* people that they must use language in only one way to try to at least make the project conceivable, but I doubt that—short of indefinite powers and a coercive will—s/he would succeed.

Compare the case of colours, where our use of them is more nearly 'devoted' and not particularly subject to will, so creativity is not much of an issue. Even here, while moving outside the head can provide interesting correlations between inner interfaces and outer properties (typically, surface spectral reflectance values) *if* one makes some very strong assumptions about the relevant environment (about which we at least know something in this case), the specific conditions at a time (such as illumination values over a range), the scale of the distal cause, the kind of medium, the degree of variation among distal causes, the specific category of cause (reflection, diffraction, illuminant), plus the state of adaptation of the organism's receptive

devices, the fatigue of the organism and its attention, the degree of approximation to something like 'normalcy' in receptive devices, the age of the organism, the distribution of receptive devices, the orientation of the receptive devices, and so on, we still have no complete theory of the nature of the 'hooks' that tie colours as experienced to distal properties. But we are a lot closer than we will *ever* be with language. With language, because of variability in human projects and the creative aspect of language use, it is not just a matter of being unworkable now, but of being impossible.

That said, I emphasize that I am sympathetic to much of the theoretical work done by Larson and Ludlow and those other formal semanticists who adopt Chomsky's LFs, such as Higginbotham, May, and Pietroski. I can say this because I think that what they in fact accomplish has nothing to do with constructing a computational-causal theory of the nature of the link between LFs and some domain outside the head. Rather, they further clarify the nature of LFs, particularly—in Larson and Ludlow (1993)—those involving the difficult propositional attitude verbs. But seen this way, their work is broad syntax, pure and simple. Indeed, one of Larson and Ludlow's claims about what to do where one *cannot* interpret in good Davidsonian fashion invites thinking of their project in these terms alone. They tackle the problem of speaking to a knotty issue for the Fregean tradition within which they work: what is an object of belief, and how does one specify what someone believes when his or her beliefs are uninterpretable according to the rules of the Davidsonian Fregean interpretation game (that is, where no extension and truth value—as these are conceived within the game—can be assigned to the words that express a person's belief)? They hold that where an assignment *can* be made the object of belief is an 'interpreted logical form' (ILF). Represent this, as above, as an <LF, I>. 'Interpretation' of a sentence is, as above, a matter of taking (the LF denoted by?) a fully specified phrase marker for the sentence and then mapping elements of an LF onto things and onto truth values.[33] Where there is an interpretation like this for a that-clause complement under a verb of belief, one can say we are told that the belief sentence says of a person that s/he believes an ILF, an <LF, I>. What, however, is the object of belief where no assignment that respects what the Davidsonian takes to be existing things and genuine truth values is possible? Here, we are told, the object of belief can be identified with a formal object, the LF side of a linguistic expression alone. For instance, if I believe that Pegasus flew (which, as it happens, I do), the that-clause of the sentence 'I believe that Pegasus flew' cannot be interpreted in the Larson and Ludlow

[33] At least, I assume that they have in mind assigning things (and classes of things) in the world to NPs, honest truth values to sentences and clauses, and so on. As I explain below, it is possible to treat interpretation as a fully formal exercise within a computational theory that stays in the head: just keep assignments of 'objects' to expressions of various sorts completely within the control of the structure and texture of LFs. If one does this, 'semantic values' become just another class of syntactic entities, and we have an honest theory again.

truth-and-reference sense, because 'Pegasus' is on this view an 'empty name'. In such a case, they say, the object of my belief is just the uninterpreted phrase-marker for 'Pegasus flew' or rather, I presume, its LF denotation; the object of my belief is a formal (syntactically defined) object alone—an LF. In my seat-of-the-pants terminology, it is Pegasus flew.

It seems to me that if Larson and Ludlow allow an uninterpreted LF as an object of belief, they might be convinced to think of LFs even without their Fregean semantic values as semantically charged entities. If they were to think of them this way, they might then think of their discussion of belief as contributing in the way I suggested to describing and explaining what the *syntax* of belief-sentences is (where syntax includes, of course, both structure and texture). That is, they could think of their work as a contribution to the lexical entry for 'believe' (its formal and semantic features), and to entries for other such verbs.

In addition, Larson and Ludlow and other like-minded semanticians could even have an honest formal theory of interpretation that relates LFs to something closer to semantic values as they conceive them and yields <LF, I>s; they only need think of semantic values differently. In what I suspect is a mildly tongue-in-cheek way, Chomsky proposes (1993a; cf. 1982, p. 324) extending the bounds of broad syntax even further beyond morphology, phonology, phonetics, and lexicology to include interpretation of a sort, and semantic values of a sort. To do this, think of interpretation as assigning to relevant parts of LFs not real things in the world, whatever they might be, but items that by stipulation have whatever features an LF contains. Think, if you will, of assigning to NPs stipulated things and classes of stipulated things, assigning to sentences (Ss) situations, and so on. Chomsky introduces a 'Relation *R*' ('for which read reference', but without the idea that reference relates an LF to something 'out there') that stands between elements of an LF and these stipulated semantic values that serve to 'interpret' it. This relation places *both* terms of Relation *R*, LFs and their semantic values, entirely within the domain of syntax, broadly conceived; they and Relation *R* satisfy my locality, intrinsicality, and mentality conditions on syntax. They are in the head. Intuitively, think of assignments ('values') as locally controlled by the formal and semantic features of relevant nodes themselves. Appealing to my decorated tree again, think of the hooks as provided by simply declaring that the 'values' suit whatever intrinsic content an LF has. This is a special form of constructivist interpretation: make your compliant world (model) by declaring it. This might even be useful. You can construct alternative models for ambiguous words, and perhaps extend it to the representation of discourse.

It is difficult, however, to know what to do with truth-values. There are two ways to conceive of the point of this effort to include interpretation in syntax, and on neither do the 'genuine' truth values of the Fregean play any role. One is to see it as an effort to show that talk of truth conditions and knowledge of truth conditions, so popular among Davidsonians and other Fregeans, can be harmlessly diverted from Davidsonian cognitive assumptions and turned to serve conceptualist-constructivist purposes. To the incan-

tation that knowing the meaning of an expression is knowing its truth conditions one can say, yes, of course. To know the meaning of an expression is to know which situation (S-node LF) has to 'obtain' for the expression to be true. But knowing this is just a matter of knowing which sentence is in question, for to know which situation it is is to know the situation the sentence stands in Relation *R* to and—assuming that the relevant sentential expression is well computed in a person's I-language—a person knows that by producing the relevant sentential expression itself, <PHON, SEM> pair that it is. Situations (after deflation) are identical with fully specified S (or IP) nodes; thus, syntax (broadly conceived) does all the work of individuating the 'truth condition', and syntactic knowledge, embodied in that competence we call the language faculty, is sufficient for knowing what the truth condition of a sentence is. These are surely trivial truth conditions: they do not invoke what people *do* take to be true, nor beliefs. Pointing this out puts the ball back in the Davidsonian's court: if you think that you can do more than this to provide a serious theory of interpretation, show us what it is. Davidson himself (1986) seems to have given up on the project.

A second way to conceive of the point of the exercise is to conceive of it as an effort to express intrinsic content—to develop a vocabulary for denoting LFs. If that is what it is, though, I think that there are better ways. Assume that interpretations consist of assigning objects and truth values, however the assignment is conceived. The problem is that mention of objects and truth values at all gives most philosophers (and particularly those inclined in Davidsonian cognitive directions anyway) the appearance of moving outside the expression to 'independently existing' things and to 'states of affairs' that make our beliefs and sentences true. Why introduce red herrings when there are perfectly good ways of expressing intrinsic content without them? I showed above how to express intrinsic content without appeal to such interpretations. We express intrinsic content directly simply by denoting it with full-fledged SEM mental event sortals that denote SEMs, supplemented by auto-exemplification where needed. Within the full deflationary project that I suggested adopting, this procedure does introduce 'objects' of a sort: situations, properties, individual intensions, and the like—without, however, the appearance that these are somehow located outside the expression, or outside the head; after deflation, these objects are various parts of LFs.

It is unlikely that Davidsonians, probably including Larson and Ludlow, will appreciate, much less adopt, this novel formal theory of interpretation, despite its real advantage of being an honest *theory* of interpretation, and available right now, for it is a theory only because it never leaves the domain of broad syntax. There is, however, an alternative. They could salvage at least some of the force of the idea that interpretation assigns things outside the head to SEMs by adopting a syntactic theory of meaning for SEMs, and glossing interpretation in terms of a highly idealized and normatively governed form of description of how people sometimes use language: objects might become referents of people's uses of terms, and truth values special forms of epistemic commitment that people undertake. This would move a major part of their

work into syntax, but put the rest into pragmatics, eliminating semantics thought of as a two-term relationship between SEMs and things.[34]

4.2 Pragmatic Interpretation

To highlight the central themes of a large number of philosophical works on meaning and content, I introduce a character I call 'Generic Wittgenstein'. Generic probably does not know anything about SEMs, or perhaps knows a little, and thinks that SEMs can be ignored, for they have no bearing on what Generic takes to be the *real* issue—that of relating language to world (interpreting). Generic typically speaks of language as being made up of 'words' and 'sentences', and thinks that the only interesting study of words focuses on how they are used. Words and sentences are 'mere tokens' in language games or something like that—perhaps vehicles for or 'bearers of sense', for example. I propose looking at what Generic has to offer by way of an account of interpretation; it is basically pragmatic description. Then, in the following section, I reintroduce SEMs and suggest some ways in which pragmatic interpretation could be improved.

Generic Wittgenstein incorporates the assumptions of those who purport to 'do semantics' by appeal to non-trivial truth conditions, uses, or conceptual roles. Collapsing these approaches into a single general approach is not, I think, irresponsible. Those who pursue them share many assumptions, including the view that the subject matter of semantics is the relationship between language and world, where this is variously construed as a matter of dealing with belief, intention and intentionality, reference, and truth (or its Sellarsian and Dummettian homologue assertibility). And at least one philosopher avowed all three approaches together: Wilfrid Sellars virtually initiated the conceptual role account, happily constructed a primitive version of what now goes under the hopeful title of a computational theory of interpretation, was an admitted truth (or assertibility) conditionalist, and expressed his debts to the spiritual founder of use theory, Wittgenstein, whose views he honoured with several articles (including Sellars, 1974). I am not saying that there were and are no disagreements among those who adopt these various strategies, but so far as I can see, all the disagreements among generic Wittgensteinians depend on different views of where to begin to describe the complex and extremely varied set of activities that are language use—intentions, beliefs, information, speech acts, judgment, truth or assertibility, reference . . .

I assume that anyone who speaks to the issue of language use must include in his or her account *at least* speakers (with intentions, desires, thoughts, needs, projects, etc., plus ears, mouth, etc.), hearers (similarly

[34] There is also, in principle, still another alternative. One could try to redefine interpretation in terms of processes in a 'performance system' on the other side of SEM/LF. That would make interpretation internalist still. Offhand, though, I doubt that what one would get would look at all like a Davidsonian/Fregean story.

equipped), things in the world (with their properties and relations), and words. Generic characteristically thinks of speakers and hearers as engaged on *tasks* or *projects* that relate them to the things of the world, and words as playing a role in carrying out these tasks. Practitioners differ in how they conceive of this task or tasks. Wittgenstein himself liberally allowed for all sorts of tasks: speakers and hearers play any number of games in which words figure. For those who adopt the conceptual role approach, however, speakers and hearers are primarily more-or-less systematic builders of theories of the world, and they focus on how words figure in this project. For the (non-trivial) truth-conditionalist, speakers are devoted truth-tellers; they focus on how words figure in truth-telling, their contribution to correct judgment.

The concept of a task that words help fulfil puts expressions and the people who use them in the domain of human action. This helps one see what kinds of theories the generic Wittgensteinians try to offer. Consider that the projects and tasks in which expressions figure include small projects and global ones—scaled all the way from Wittgenstein's builder's language games to truth-telling in general. But in all cases the primary measure of words is taken to lie in the extent to which they lead to success in the project—the way in which they advance the aim of the action(s) in which the words are involved. For the truth-teller, a sentence is seen as a judgment that is measured by the extent to which it is correct (or at least reliable). For the builder, the word is measured by the extent to which it leads to a successful outcome: the beam is carried, and it is properly placed. Call the value assigned to a 'word' in this way a *use value*. A use value can be assigned by appealing to any number of ways of succeeding, or partially succeeding, at some task in which words figure; what counts as success depends on the task. Examples include truth, correctness, plausibility, appropriateness, politeness, describing (a success term), gathering information (another), referring (still another), and perhaps even intending to refer. Looked at in this way, it is clear that the truth-conditional, use, and conceptual role 'theories of meaning' that Generic offers—in effect, virtually all those 'theories of meaning' that philosophers have taken seriously for the past 50 years or so—are normative through and through.

Wittgenstein himself thought that the combination of an enormous domain, an unlimited number of tasks, and a normative dimension made it impossible to construct serious theories of the use of words. Language use is too diverse—we play all kinds of games with language, perform all kinds of tasks—too vague, and too context-sensitive to yield even a uniform classification of uses or roles, much less an explanatory theory that routinely connects word, person, and world—a sort of computational theory. And he emphasizes, if anyone does, that language use is normatively governed. It is subject to criticism and assessment by 'masters' of a practice; it is done well or ill, appropriately or not. All this, he seems to believe, makes the prospects of serious theorizing dim; he treats the desire for theory in this domain as a disease, and offers therapy. Clearly, his observations present

Chomsky's creative aspect of language use from another point of view—without in Wittgenstein's case any idea at all of how a serious theory of language could be constructed.

These admonitions do not in themselves rule out some modest 'theories' that are no more than systematic attempts to classify words in terms of their uses. Where Wittgenstein restricted himself to such context-sensitive descriptive devices as perspicuous representations and language games, Austin offered a more general classification of ways in which language is used under the rubric of speech acts. Another descriptivist, Searle, has tried to turn this kind of classification into lessons about intentionality; but while this might place a classification of what we can do with words in another context and help serve a philosophical end, it is not a serious theory. It attempts to 'explain' in those ways Sellars might have called showing people a way around in the philosophical neighbourhood; it does not turn classification into a serious theory of use. To do that, many Generics have thought, we need causal explanation, probably in computational form.

Generics who hope to move beyond 'mere description' to systematic generalization try to construe language users as input–output devices, information-theoretic engines, embodied theories of the world, truth-telling engines, and the like: speakers are gatherers of truths, perhaps so that they can make good or reliable decisions, or perhaps because that is just what it is to be a speaking human. Thus, Sellars's functional classification of words is accompanied by a story about how we all play a uniform language game consisting of 'language entry rules' (perception), 'language transition rules' (inference), and 'language exit rules' (acts/action). His and related efforts at turning classification by use values to the purposes of explanation are attempts to convince us that measuring words by informational 'content' or by truth-in-a-theory somehow converts a descriptive exercise into a causal theory of how the mind/brain functions in the world. We are offered a picture in which words have the characteristics they do (informational contents, truth-values) when the machinery is working properly (another success term). Unsurprisingly, the machinery is rarely specified, except by gesturing in the direction of neurons and firing rates, or photons and retinas. But these are only gestures: until put in a serious proposal, there is no way to judge whether they are on the right track, or where to look for a better view. Like Wittgenstein, and from another perspective Chomsky, I am very suspicious of a claim from someone that they have a theory in this domain, and again like them, for reasons already given, I think it is extremely unlikely that a theory can be constructed in the domain.

Someone will object that it *must* be possible to construct a theory in this domain. We could not communicate or understand one another unless language use were regular and even predictable. If it is regular and predictable, we can construct a science, can we not? That language use is (often) regular and predictable does not show that there must be a science of use. No one denies that we routinely assign use values to words on occasions, given knowledge of person, task, circumstance, and so on. In fact, familiar with

the various tasks we perform with language as we are, sensitive to what makes a difference in them, and reasonable judges as we are about how reliable (or inclined or not to whimsy, caprice, irony, distraction, boredom . . .) a person is, we can not only assign use values, but often predict what a person will say on an occasion, or—adapting Wittgenstein—what path(s) s/he will follow, given that s/he is playing a particular game. This fact, however important it may be in coming to translate what another says, or interpret another, should not be construed as evidence that this is a domain in which one can construct a serious theory. Our capacity to assign use values, given what we know about people, knowing how to make adjustments in judgment (without being able to say ahead of time what might be required), realizing what contextual factors might be relevant, cannot be captured by anything that looks like a theory. Basically, we have no way to state boundary conditions. We cannot save the day by calling on *ceteris paribus* clauses, or the like. All *what* things equal?—well, it depends on the task the person (and whether s/he is reliable, whimsical . . . at certain times of day/all the time/when dealing with Gerald, but not Harriet . . .), the context, the . . . and there is no way to limit any of this or anticipate what will be needed—much less turn it into a science. Keep in mind that we come up with assignments of use values even in spectacular cases, as Davidson's Mrs Malaprop illustrates. But we follow no single algorithm or discernible finite list of them in doing so, whether in Mrs Malaprop cases or others.

Basically, human action, though amenable to description—sometimes at least insightful—is in no serious sense amenable to scientific explanation. Action is driven by all sorts of aims, and fulfils all sorts of ends. We are not dealing with automatons or faculties like colour vision or the language faculty, which are insensitive to correct and incorrect, appropriate and inappropriate. We are dealing with variably rational agents engaged on various tasks, certainly not just one. Perhaps it is possible to talk seriously of economic rationality; it is unlikely that one can talk seriously of rationality in language use. Perhaps there is rationality of a systematic sort in language use in barn-building, and some other tasks; but is there rationality of a systematic sort in language use in poem-writing, or even speech-giving? All this reflects, again, Descartes's observation, revitalized by Chomsky. Human language use is creative. Although as used by a rational agent a word or sentence is usually appropriate to circumstance and we can judge that it is (we can assign a use-value to it or think it appropriate to circumstance, with more or less confidence), nevertheless, language use is unbounded and stimulus-free. That is, we can—by taking into account a person's desires and interests—see why a word is appropriate for some circumstance; this is providing a Wittgensteinian 'path that someone might follow'. But it is very far from providing a causal or computational theory of the relationship between things in the world, sense organs, brain, and outputs. It is no theory of behaviour. None of this, of course, precludes causal theories of language. The computational theory that produces PHONs and SEMs is one. But that, clearly, is not a theory of language use; its domain is totally inside the head,

and it seems it must be, as a condition for having a theory of language at all. So it appears that we can have a naturalistic theory of meaning if it is understood as a theory of SEMs, but no naturalistic theory of language use. Wittgenstein was right to stick with classification in this domain. And Chomsky is right to think that a science of mind (the language faculty) is possible only if it is internal.

4.3 Could SEMs Guide Use?

Calling the internalist domain over which SEMs are defined the domain of *meaning*, I propose thinking of the pragmatic domain where use values are assigned to words the domain of the *meaningful*.[35] This terminology attempts to capture the idea that use values speak to the ways in which words 'matter' to us—the ways in which they help or hinder the fulfilment of the various projects in which they figure. This terminology recalls C. I. Lewis's usage; the aim of an account of the meaningful is to get a grip on why a word counts, or is important to us. This is also, I suspect, what most people are asking for when they ask what someone means or intends when s/he says something.

Prospects for a theory of the meaningful are dim, we found, but that was on the assumption that we are dealing with words, not SEMs. Would prospects change, and perhaps improve, if an internalist SEM theory of meaning were somehow thrown into the hopper? There is an interesting way to conceive of the effect of SEMs on pragmatic efforts at interpretation which takes a Wittgensteinian analogy and turns it in a different direction. Wittgenstein recommended looking at words as tools that we use to further our projects. I have no doubt that he thought that describing tools is exhausted by detailing their function(s)—adapting earlier terminology, saying (by

---

[35] Notice that the division I propose between the domain of meaning and the domain of the meaningful is not the distinction that those who speak of a 'dual aspect' semantics have insisted on. 'Dual aspect' semantic theories split a theory of what they choose to call 'meaning' into a conceptual role theory of understanding and a truth conditional theory that relates language to world, but they hold firmly to the principle that there is a gulf between syntax and semantics, including meaning. See, for example, McGinn, 1982; Block, 1986. For them, the primary aim is to keep apart the supposedly separate contributions of truth conditions and conceptual roles to a combined theory of meaning, assigning (usually) a conceptual role theory the task of dealing with *understanding* meanings (presumably because conceptual roles look mentalistic) and a truth conditional theory the task of dealing with the relationship of a language to the world.

Incidentally, if the basic split is between syntax and pragmatics, it is reasonable to introduce two concepts of understanding for the two domains, one computationally characterized and associated with meaning, and the other dealing with various aspects of language use—capturable in different but complementary and not exclusive ways by conceptual roles, truth conditions, language games, etc. (See McGilvray, 1991 for a related view.) Bluntly, understanding meanings is a matter of having a syntactic competence and the rich store of tacit knowledge that goes with it; understanding use is a matter of being able to make correct judgements where this ability is capturable in various ways. The second *ability* presupposes the first *competence*.

describing) what their use values are. But it is possible to look at tools not only in terms of the ways they are used (their functions) but in terms of their intrinsic characters. Consider a hammer. Often the first thing that comes to mind when someone is asked what a hammer is is the use to which it is put; this is typical of artefacts. But there are also non-functional ways to classify hammers. One such way classifies (describes) them in terms of their structure and composition: they come with handles (usually of wood, plastic, or metal, coated or not) and with claw, peen, or other-shaped heads, made (usually) of moderately malleable and reasonably dense metals. Now, it is surely plausible that hammers can only perform the tasks that they do—the tasks that we are interested in having them accomplish—by having certain structures and by being made of certain materials. Thus, one might say that their structure and compositional texture constrain and guide their use, although they do not fix their use determinately. I can use a claw hammer to prop up a table by placing it beneath one of a table's unhappily shortened legs. But while a possible use, and in some sense allowed by the hammer's structure and texture, this is not a suitable use for a claw hammer: given what it is, it serves other tasks much better.

I suggest that words are similar. They can be classified not just in terms of their use values but, thought of as SEMs or parts of SEMs, one can say what their structure and texture is in a very detailed way. Pursuing the analogy, think of the syntactic description of SEMs as a description in a sort of materials science, one that says what the tool is made up of. Continuing, perhaps the material out of which a SEM is composed makes it more suitable for some uses than others. If so, perhaps there is room for another way to look at the domain of the meaningful. Specifically, perhaps one can find in SEMs themselves some clues to the most puzzling aspect of the creative aspect of language use, that while stimulus free and unbounded, linguistic expressions are also (typically) appropriate to the circumstances in which they are used.

First, a caveat: the analogy between expressions and artefacts should not be pressed too far, for there are several crucial differences between expressions and artefacts. I mention one important one. Hammers, like other artefacts, are specifically designed and manufactured by humans to serve certain purposes. Recognition of these purposes has led craftsmen and engineers to use certain kinds of materials in their manufacture; as a result, there is an acknowledged, function-driven reason that hammers have the structure and compositional texture they do, and for specific purposes, hammers made from different materials can be produced. Hammers are 'cognitively penetrable', and their structure and texture is put into them by engineers to enable them to serve their statable functions. In the case of lexical items, however, there are no engineers; one can only combine given items in different ways (Saussurean arbitrariness). One can perhaps think of what is fixed for phonetic, formal, and semantic features (the first two with parameterization) as the result of evolutionary development over centuries by the human species interacting with its environment, although I doubt it

is possible to sustain a selectional hypothesis that shows this. However, this undoubtedly right but unprovable evolutionary claim does not support the idea of a cosmic engineer, and it certainly does not support social and individual engineering of SEM structures and textures.[36] The capacity that each of us has that is described by a computational syntactic theory is largely built in at birth, and cannot be re-engineered. The range of intrinsic contents available to all humans is not only relatively uniform, but it does not really change: it is given to us, not 'learned'. So the characters of the natural language tools we have available to us are fixed, not plastic.[37] Extraordinarily complex SEMs can be constructed from limited resources, but that is not the kind of plasticity I have in mind. The possible structures and textures of the SEMs of an I-language are not put in our heads by any one of us, nor are they produced by any generalized theory-construction capacity that we might or might not have. They are not the products of communities—not due to the anxious training of our parents, nor to the adjudicative efforts of a 'linguistic community' that wants us to conform to epistemic, political, and social principles of correct usage. Engineering modifies use alone—and that only within limits—not intrinsic contents. But there is no loss in any of this. First, an infinite number of SEMs is available, including an infinity that have not been used; the expressive power of natural languages cannot be exhausted. Second, the gains for language acquisition are extraordinary; fixity is a *sine qua non* for acquiring a language at all. And third, as Cudworth and Herbert

---

[36] Remember that there is Saussurean arbitrariness, and that natural language lexicons can vary in scope and selection. Surely—*pace* Fodor, 1982—one can also introduce lexical items that are defined in terms of innately fixed features: 'refrigerator' is a plausible example, as is 'internet'. It is difficult to know what the limits, if any, of this kind of invention might be without a more complete theory of the lexicon.

[37] As note 36 indicates, there is a form of linguistic invention. I only deny that there are clear cases of meaning-invention (as opposed to generation) for the expressions of a *natural* language. Obviously, we can construct scientific theories, often laboriously, and the terms of these theories really do have invented meanings. But scientific theories are not natural languages. They are learned late, if at all; they satisfy none of the poverty-of-stimulus criteria that natural language expressions do.

What, however, of 'grue' (Green until 3:00 in the afternoon (GMT) on 1 January 2000, and blue thereafter') and the like? These, unlike H_2O, are not terms in a scientific theory. They are philosophers' cryptograms, collapsing complex claims about when to make assertions using various natural language terms into a form of code. No intrinsic content is defined by appeal to assertion at different times, nor by assertion at all, and it is not here; what one has here has no bearing on meaning (intrinsic content) at all. It is just a puzzle concerning justified assertion; it is another way to raise sceptical questions about what counts as meaningful, not what counts as meaning.

Similar things are true of metaphor, which is a matter of taking an expression that has been used to refer to and describe certain things, and using it in a different way—usually, it is a matter of referring to something of a different kind, as with the now almost-literal claim 'George is a wolf'. This is not invention of meaning, but different and potentially illuminating application that adds to what we take to be meaningful. So, while with scientific languages there is invention of meaning, but the language is not natural, with 'grue' and metaphor, there is no invention of meaning at all, but puzzles concerning the use of tools, and unusual exercises of tools.

of Cherbury pointed out, we could not develop perception and knowledge of the world, or even moral judgment, if we did not have internal, innate guidance. So we are well served by fixity.

Pursuing the analogy again, Cudworth and Herbert's point is particularly apropos for an effort to say how internal meanings (their 'ideas' and 'notions') might, after all, bear on use—how intrinsic contents could speak to appropriateness of use. One way to conceive of how they might is, like them, to ask to what extent some 'natural kind' terms such as 'water' with their intrinsic contents could, at the interface level (SEM), affect other cognitive faculties, such as vision and touch. This is, I think, a fruitful route to pursue, and—unlike Cudworth and Herbert's time—perhaps it is now just barely possible to follow it; but it is well beyond the scope of this paper. Another way is to think of intrinsic contents as guiding the assignment of use values to judgments made in language. Specifically, look for ways in which SEMs with their intrinsic features could guide judgment. Use values come in a number of forms, depending on the task(s) at hand: correctness of inference, perlocutionary success, politeness and felicity values, epistemic correctness and truth, and the like. Generally speaking, use values measure *appropriate* use, given the task to which a/the SEM is put. With this in mind, look at correctness of inference. I say Harriet is chasing George. If so, she is following him, and both following and chasing intentionally. I say Mort is painting his house; if so, he is painting the outside. These are 'analytic truths' (indeed, a priori analytic truths). Similarly, I say Mort has started a book. He has started to read a book, or he has started to write one. Anyone who has the relevant vocabulary items at all realizes this, and everyone relies on this kind of detailed (but implicit) knowledge of structure and texture while using these and thousands of other lexical items. Surely, then, there is a strong sense in which SEMs guide their use. It is guidance alone: they do not control it. I can say that Harriet is chasing where she is following him and is not intentionally chasing him, and perhaps even has no idea at all that she is behind him; she is just coming along after him. And I can say that Mort has started a book where he has just placed it in a wood chipper, which has begun to throw out pieces of paper.

The concept of intrinsic content in linguistic expressions guiding use, and through that the assignment of use values, points towards a different and, I think, more productive account of language use than that found in thinking of people as task-doers, using 'words' as tools that they make. The approach I have in mind begins with theories of human innate capacities and finds in what these theories say about their outputs an outline of the intrinsic but implicit knowledge a person brings to action. An approach to use of this sort relies on internal resources for which one *can* construct naturalistic theories to define, in the end, a kind of ecological niche for the human organism. Here 'explanation' in the case of language use does not consist in constructing—*per impossible*—a complete computational account of language production, given circumstances. There is, then, no science of use. It consists in a fuller elaboration of the various internally produced perspectives that

the exercise of innate capacities provides, plus an elaboration of how these interact with other systems in the head (including other cognitive faculties and what Chomsky calls 'performance systems') to speak to how the organism uses what it has available in all sorts of enterprises, from perception, to food-gathering, to writing poetry. Proceeding in this way, one sees how, given our natures, we 'make' our worlds in the way we do. A constructivist understanding of the way in which linguistic perspectives can contribute to the construction (perception and judgment) of the world, joined to the nativist and internalist theory that scientifically explains the production of these perspectives, seems to be the key. Cudworth was on the right track to speak not just of the mind 'protruding' innate ideas, but of these innate ideas constituting the 'measure of . . . things'. Chomsky put it in a different way in *Language and Mind* (p. 13), where he said that while a science of linguistic behaviour looks to be impossible, for there is no causal explanation of language use, it might be possible to use science to explain—although not causally—how 'intelligent behaviour' (of which language use is a paradigm case) is *possible*. This is a bit like Kant's effort to say how judgment of various sorts is possible. Unlike Kant, however, Chomsky has a science of the meanings that figure in judgment.

Department of Philosophy
McGill University

References

Baker, M. 1996: *The Polysynthesis Parameter*. Oxford University Press.
Bigelgac-Babic, R, Bertoncini, J. and Mehler, J. 1993: How Do 4-Day-Old Infants Categorize Multisyllabic Utterances? *Developmental Psychology*, 29, 4, pp. 711–21.
Bilgrami, A. 1992: *Belief and Meaning: The Unity and Locality of Mental Content*. Oxford: Blackwell.
Block, N. 1986: Advertisement for a Semantics for Psychology. In Uehling and Wettstein (eds), *Midwest Studies in Philosophy*, 10, 615–77.
Burge, T. 1986: Individualism and Psychology. *Philosophical Review*, 95, 3–45.
Burge, T. 1989: Wherein is Language Social? In A. George (ed), *Reflections on Chomsky*. Oxford: Blackwell, 175–91.
Chomsky, N. 1957: *Syntactic Structures*. 's-Gravenhage: Mouton.
Chomsky, N. 1959: Review of Skinner, *Verbal Behavior*. In *Language*, 35, 1. pp. 26–58.
Chomsky, N. 1965: *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
Chomsky, N. 1966: *Cartesian Linguistics*. New York: Harper and Row.
Chomsky, N. 1975: *Reflections on Language*. New York: Random House.
Chomsky, N. 1982: *Lectures on Government and Binding*. Dordrecht: Foris.
Chomsky, N. 1986: *Knowledge of Language*. New York: Praeger.
Chomsky, N. 1992a: Language and Interpretation: Philosophical Reflections and Empirical Enquiry. In J. Earman (ed.), *Inference, Explanation, and Other Philosophical Frustrations*. Berkeley: University of California Press.

Chomsky, N. 1992b: A Minimalist Program for Linguistic Theory. *MIT Occasional Papers in Linguistics,* No. 1. Also in Hale and Keyser (eds.) 1993, *The View from Building 20*. Cambridge, MA: MIT Press.
Chomsky, N. 1992c: Naturalism and Dualism in the Study of Language and Mind. Typescript of Evans Lecture. Printed in *International Journal of Philosophical Studies*, 2 (1994), pp. 181–209.
Chomsky, N. 1993a: Explaining Language Use, MS. Published in Tomberlin (ed.), *Philosophical Topics*, 20, 205–31.
Chomsky, N. 1993b: *Language and Thought*. Monograph 3 in Anshen Transdiciplinary Lectureships in Art, Science, and the Philosophy of Culture. Wakefield, RI: Moyer Bell.
Chomsky, N. 1995a: Language and Nature. *Mind*, 104, 1–61.
Chomsky, N. 1995b: *The Minimalist Project*. Cambridge, MA: MIT Press.
Chomsky, N. 1997: Internalist Explorations. MS., to appear in a volume of papers in honour of Tyler Burge, ed. B. Ramberg.
Chomsky, N. 1998: New Horizons in the Study of Language. MS.
Christophe, A., Dupoux, E., Bertoncini, J. and Mehler J. 1994: Do Infants Perceive Word Boundaries? An Empirical Study of the Bootstrapping of Lexical Acquisition. *Journal of the Acoustical Society of America*, 95, 3, pp. 1570–80.
Clark, A. 1993: *Sensory Qualities*. Oxford: Clarendon.
Cudworth, R. 1976 (1731): *A Treatise Concerning Eternal and Immutable Morality*. New York: Garland. Reprint of 1731, posthumously edited by Chandler, Bishop of Durham.
Davidson, D. 1984: *Inquiries into Truth and Interpretation*. London: Oxford University Press
Davidson, D. 1986: A Nice Derangement of Epitaphs. In LePore (ed.), *Truth and Interpretation*. Oxford: Blackwell, pp. 433–46.
Egan, F. 1995: Computation and Content. *Philosophical Review*, 104, 181–203.
Fodor, J. 1980: Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology. *Behavioral and Brain Sciences*, 3, 63–73.
Fodor, J. 1982: The Current Status of the Innateness Controversy. Repr. in Fodor, *Representations*. Cambridge, MA: MIT Press.
Fodor, J. 1983: *Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.
Fodor, J. 1987: *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press.
Fodor, J. 1994: *The Elm and the Expert*. Cambridge, MA: MIT Press.
Goodman, N. 1952: On Likeness of Meaning. In Linsky, (ed.), *Semantics and the Philosophy of Language*. Urbana: University of Illinois Press.
Goodman, N. 1961: About. *Mind*, 70, 1–24.
Goodman, N. 1968: *Languages of Art*. Indianapolis: Bobbs-Merrill.
Gleitman, L., Fisher, C., Hall, D.G. and Rakowitz, S. 1994: When it is Better to Receive than to Give: Syntactic and Conceptual Constraints on Vocabulary Growth. *Lingua*, 92, 333–75.
Grimshaw, J. 1992: *Argument Structure*. Cambridge, MA: MIT Press.
Grimshaw, J. 1994: Lexical reconciliation. *Lingua*, 92, 411–32.
Herbert of Cherbury, 1937 (1622): *De Veritate*. Tr. Carre. Bristol: University of Bristol Press.
Higginbotham, J. 1985: On Semantics. *Linguistic Inquiry*, 16, no. 4.
Hornstein, N. 1991: Grammar, Meaning, Indeterminacy. In Kasher (ed.), *The Chomskyan Turn*. Oxford: Blackwell, 104–21.

Hornstein, N. 1995: *Logical Form: From GB to Minimalism*. Oxford: Blackwell.
Larson, R. and Ludlow, P. 1993: Interpreted Logical Forms. *Synthese*, 95, 305–55.
Marr, D. 1982: *Vision*. New York: Freeman.
May, R. 1985. *Logical Form: Its Structure and Derivation*. Cambridge, MA: MIT Press.
Mehler, J. and Christophe, A. 1994: Language in the Infant's Mind. *Philosophical Transactions of the Royal Society of London Series B*, 346, 13–20.
McGilvray, J. 1983: To Color. *Synthese*, 54, 37–70.
McGilvray, J. 1991: *Tense, Reference, and Worldmaking*. Montreal and Kingston: McGill-Queen's.
McGilvray, J. 1994: Constant Colors in the Head. *Synthese*, 100, 197–239.
McGinn, C. 1982: The Structure of Content. In A. Woodfield, (ed.), *Thought and Content*. Oxford: Clarendon.
Moravcsik, J. 1975: *Aitia* as Generative Factor in Aristotle's Philosophy. *Dialogue*, 14, 622–36.
Moravcsik, J. 1990: *Thought and Language*. London: Routledge.
Pietroski, P. 1994: A Defense of 'Derangements'. *Canadian Journal of Philosophy*, 24, 95–117.
Pinker, S. 1989: *Learnability and Cognition*. Cambridge, MA: MIT Press.
Pinker, S. and Bloom, P. 1990: Natural Language and Natural Selection. *Behavioral and Brain Sciences*, 13, 707–84.
Pinker, S. 1994a: How Could a Child use Verb Syntax to Learn Verb Semantics? *Lingua*, 92, 377–410.
Pinker, S. 1994b: *The Language Instinct*. New York: William Morrow.
Putham, H. 1975: The Meaning of 'Meaning'. In H. Putham, *Mind, Language, and Reality*. Cambridge University Press.
Segal, G. 1991: Defence of a Reasonable Individualism. *Mind*, 100, 485–93.
Sellars, W. 1963: Philosophy and the Scientific Image of Man. *Science, Perception and Reality*, ch. 1. London: Humanities Press.
Sellars, W. 1974: Meaning as Functional Classification. *Synthese*, 27, 417–37.
Wittgenstein, L. 1953: *Philosophical Investigations*. New York: Macmillan.