Conjoining Meanings:

Semantics Without Truth Values

Paul M. Pietroski
University of Maryland
(DRAFT: 21 Sept 2014)
Chapter Zero: Getting Started

Many animals undergo dramatic changes after birth. Caterpillars go through a pupal stage and emerge with scaled wings. Young humans acquire words, which can be combined to form boundlessly many expressions that are both meaningful and pronounceable. Such changes can transform an animal’s lifestyle. Butterflies flutter. People talk. This book is about linguistic expressions, their meanings, and how these meanings are related to human nature.

Following Chomsky (1957, 1964, 1965, 1977, 1986, 1995), I think human languages—the spoken or signed languages that children naturally acquire given ordinary experience—are biologically implementable procedures that generate expressions that connect meanings with pronunciations. My proposal is that these meanings are instructions for how to build mental symbols of a special sort. More specifically, I think lexical meanings are instructions for how to access concepts of a special sort, and phrasal meanings are instructions for how to build mental predicates (a.k.a. monadic concepts) that involve a distinctive kind of conjunction. On this view, the meaning of ‘brown cow’ is an instruction, and executing it yields a complex concept that can be used to classify something as both a cow and a brown one. More generally, I claim, human languages are procedures that generate pronounceable recipes for making conjunctive mental predicates. Meanings are the recipes. This hypothesis has to be explained and elaborated before it can be defended and compared with alternatives. But the first thing that needs explaining is the topic. What am I talking about when I talk about meanings?

1. The Topic

If a biologist says that butterflies are mature insects that used to be caterpillars, one can know what she is talking about yet be surprised by the claim. A person can know what butterflies are without knowing much about their nature or history; and a biologist can safely assume that most people know what butterflies are in this minimal sense. But if an astronomer says that quasars are unusual galaxies, one might ask what a quasar is. Similarly, if a linguist or philosopher says that meanings are instructions for how to build mental symbols, one might not know what he is talking about, even given tolerably clear conceptions of mental symbols and how complex things—like cakes, strings of numerals, and proteins—can be assembled via processes that can be described as executions of instructions. Pointing to meanings is harder, or perhaps less helpful, than pointing to butterflies. Moreover, specialists don’t even agree about whether there are meanings that linguistic expressions have.¹ So we can’t exclude the possibility that every proposal of the form “Meanings are Xs” is wrong. Yet to evaluate such proposals, we need a working conception of what meanings are if there are any. So before turning to my own proposal in chapter one, I want to articulate a working conception of meaning, in part by reviewing some basic facts and familiar claims.

1.1 Human Languages: Unbounded yet Constrained

It can be tempting to start with claims about how meaning is related to communication, truth, intention, reference, inference, or necessity. For example, one might assume that the meaning of ‘cow’ somehow maps conversational situations onto sets of cows, or that ‘cow’ is understood to be true of cows. I think that these particular assumptions are false, and that even as idealizations, they impede discovery of how linguistic expressions can be used to communicate intentions, refer to things, and make truth evaluable claims that have logical implications. But in any case, for purposes of initially characterizing meaning(s), I think the natural phenomenon of human

¹ Some think that talk of meaning is, at best, talk of certain ways in which expressions can be equivalent; see, e.g., Quine (1960). Compare talk of number, about which there has also been disagreement.
language acquisition provides a better anchor for discussion. Any starting point must ultimately be judged in terms of where it leads. Though we can try to begin uncontrovertially.

Unlike a caterpillar, a human infant can become a speaker of English, given an ordinary course of experience and development. Of course, acquiring English as opposed to Japanese requires an environment of a certain sort. But caterpillars cannot acquire either language in any setting; and likewise, it seems, for ravens and orangutans. Humans are special in this respect. Human languages are also distinctive. These “child-acquirable” languages differ, in various ways, from non-human communication systems (e.g., bee dance) and the formal systems that adult theorists have explicitly invented. One striking feature of human languages is that they exhibit ambiguity in an interestingly constrained way.

To illustrate with a famous kind of example, (1) is a string of words that speakers of English can understand as a sentence that is roughly synonymous with (1a) as opposed to (1b).

(1) the guest is easy to please
(1a) It is easy for us to please the guest.
(1b) #It is easy for the guest to please us.

The symbol ‘#’ indicates that (1) cannot be understood as a sentence that can be paraphrased with (1b), even given loose standards of paraphrase. If (1) is understood as an expression of English, rather than a mere string of words, it is understood along the lines of (1a) and not (1b). By contrast, string (2) can be understood as a sentence that is roughly synonymous with (2b), but not as a sentence that could be paraphrased with (2a).

(2) the guest is eager to please
(2a) #The guest is eager that we please her.
(2b) The guest is eager that she please us.

Yet (3) is ambiguous; it can be understood along the lines of (3a), or along the lines of (3b).

(3) the guest is ready to please
(3a) The guest is ready to be pleased.
(3b) The guest is ready to be a pleaser.

Later, I’ll need to say more about sentences, indicated above with capitalization and periods. But for now, let’s make do with the old idea that sentences are pronounceable analogs of “complete thoughts.” The important point is that a string of words can be interestingly unambiguous. Competent speakers of English can recognize that (3) has two meanings—in the sense of being understandable in two ways—while the first two strings have one meaning each.

This is an initial gesture at the meanings I am talking about: string (3) has more than one. In this respect, (3) is like the pronunciation ber. Using traditional terminology, ‘bear’ and ‘bear’ are homophones; and spoken English presumably has at least two words that get spelled the

---

2 See Chomsky (1964). Austen (1813) describes a scene in which “Bingley was ready, Georgiana was eager, and Darcy determined to be pleased.” This description is unambiguous, even if each character was also ready, eager, and determined to please. If it helps, compare (3) with ‘the duck is ready to eat’. Each of (1-3) also fails to have the meaning that ‘elephants arrived’ has. But this absence of ambiguity is relatively uninteresting.

3 This does not imply that speakers are infallible. It is conceivable, though unlikely given other examples, that (3) is univocal but “made true” in distinct ways; compare ‘the guest entertained us’. Even if (3) is ambiguous, one might try eschew meanings by saying that (3) is equivalent to (3a) in a relevant respect R and equivalent to (3b) in another relevant respect R*; (1) is R-ishly equivalent to (1a) though not R*-ishly equivalent to (1b); and (2) is R*-ishly equivalent to (2b) though not R-ishly equivalent to (2a). This proposal—concerning what it is for (3) to “have the meanings” it does have, and for (1) and (2) to “not have the meanings” they don’t have—implies that the relevant equivalence classes are not determined by the meanings of expressions. So one wants to see meaning-free specifications of the alleged dimensions of equivalence, given a range of relevant examples, including (4-20) below.
second way. Speakers of English somehow connect several meanings with ber, which is thereby connected with several meanings. Let’s not worry here about what pronunciations are, or how to individuate them. But a soprano and a bass can pronounce ‘bear’ in a same way, as can children who articulate in various ways. My suspicion is that pronunciations are instructions geared to articulatory/perceptual systems; see Halle (1990), Chomsky (1995). Though for purposes of initially characterizing meanings, any plausible conception of pronunciations will do.

As lexicographers know, it can be unclear where homophony ends and polysemy begins. But there are many clear cases of the former. The pronunciation of (4) has several meanings

(4) the sheriff near the bank is eager to draw

because the pronunciations of ‘bank’ and ‘draw’ each have more than one. In this respect, (3) is different. The pronunciation of (3) has two meanings, but not because any word in (3) is a homophone. Put another way, the ambiguity of (3) does not depend on lexical homophony.

Here and throughout, I use ‘lexical item’ to talk about (meaningful) expressions of a human language that have no meaningful constituents. So a word need not be a lexical item. The word ‘pleased’ is not a one, given the tense morpheme ‘d’, and likewise for the plural noun ‘cows’. However, ‘pants’ may be a lexical item. Identifying the lexical items of a human language takes work; see, e.g., Borer (2005). But there is at least one with the pronunciation kao, even if the singular noun ‘cow’ is the result of combining a lexical root with a meaningful count-noun feature. And the lexical item ‘ant’ is not a constituent of ‘pants’. This much is obvious. The lexical item ‘saw’ that is used to talk about saws is homophonic with the past tense of ‘see’. But we needn’t—and shouldn’t—decide in advance whether the verb used to talk about using the tool is itself a lexical item. For purposes of indicating what meanings are, it doesn’t matter if most words turn out to be complex expressions. Whatever lexical items are, we need to distinguish lexical homophony from the combinatorial or structural homophony illustrated with (3).

If only for simplicity, let’s assume that each lexical item connects one meaning with one pronunciation, allowing for null pronunciations as special cases. (I return to the costs of rejecting this assumption.) However we count lexical items, a typical child acquires thousands. But since each atomic meaning-pronunciation pair must be encoded in some retrievable form, each human language has finitely many lexical items. So while English offers many examples of lexical homophony, they can be listed. On the other hand, there are unboundedly many examples of combinatorial homophony. Note that (5) is like (3) in having two meanings.

(5) ducks that are ready to eat with a fork are eager to eat with a spoon

There are also unboundedly many examples, including (6), of interesting non-homophony.

(6) ducks that are ready to eat with a fork are eager to eat with a spoon

As often remarked, lexical items can be combined to form boundlessly many expressions. But whatever expressions are, the “generative capacity” of a human language is unbounded and constrained in ways that presumably reflect the biologically implemented human capacities to generate expressions that are meaningful and pronounceable. Each (non-null) pronunciation $\pi$ of

---

4 Consider ‘I had to saw the wood without a saw’. One complication is that irregularity, as with the past tense of ‘see’ or ‘bring’, may lead children to acquire lexical items whose meanings have the meanings of other lexical items as constituents. For example, ‘brought’ may be a lexical item whose meaning can be analyzed in terms of ‘bring’. More generally, the lexicon of a human language may exhibit certain redundancies; see Pinker (1988) for relevant discussion. But I suspect that, even among human linguistic expression, atomic or complex, has a meaning—even if there are unpronounced expressions that can be the grammatical subject or object of ‘please’ in (1–3). I will come back to the potential relevance of this asymmetry, and whether it matters if we describe lexical items as connecting meanings with pronunciations or pronunciations with meanings; cp. Chomsky (2000a).
a human language $H$ has a certain number of meanings in $H$. It may be vague whether $\pi$ has four or five meanings. But modulo vagueness, $\pi$ will have $n$ but not $n+1$ meanings, for some $n$; and in many cases where $\pi$ is the pronunciation of a word-string, $\pi$ will have fewer meanings than the pronunciation of a superficially similar word-string, even controlling for lexical homophony.

Constraints on combinatorial homophony are often discussed in the context of language acquisition, because of their implications for the innate endowment that allows children to project from somewhat idiosyncratic courses of experience to the languages actually acquired, as opposed to superficially similar languages that exhibit even more (or still less) homophony. But whatever one says about the relative contributions of nature and nurture to acquisition, one can begin to get a fix on what meanings are, by attending to what pronunciations can and can’t mean.

Endlessly many English pronunciations have zero meanings, even restricting attention to pronunciations of strings of English words, as opposed to pronounceable but meaningless strings like ‘flib gork’. For example, (7-11) below are examples of word salad, even though (12) is easily understood as a fine sentence; see Chomsky (1957).

(7) *we been have might there
(8) *we been might have there
(9) *we have been might there
(10) *we have might been there
(11) *we might been have there
(12) we might have been there

The asterisks indicate that (7-11) are not understood as acceptable English sentences. So in particular, these strings are not understood as acceptable variants of (12). But the key fact regarding (7) could be reported less economically with (13).

(13) we been have might there
(13a) #We might have been there.

When a child acquires English, she acquires a language that connects (7-11) with zero meanings; see Berwick et. al. (2011) for discussion. Likewise, endlessly many strings have exactly one meaning. It may be unsurprising that (12) has one meaning and not two. But as (6) illustrates, it is not trivial that given a course of experience which in fact leads a child to acquire English, the child acquires a language in which combinatorial homophony is as constrained as it is in English.

Constraints can manifest in ways that are subtle though clear upon reflection. Note that string (14) has three meanings, even with ‘saw’ understood as the past tense of ‘see’.

(14) we saw the spy walking towards the store
(14a) We saw the spy while we were walking towards the store.
(14b) We saw the spy who was walking towards the store.
(14c) We saw the spy walk towards the store.

Unlike (14c), (14b) can be used to describe a situation in which we saw the spy without seeing him walk; see Chomsky (1964, p.73). But (15) has only the interpretation corresponding to (14c).

(15) this is the store that we saw the spy walking towards

So (15) is interestingly unambiguous.

---


6 I’ll return to “perceptual reports,” which are independently interesting. But note that if ‘saw’ is understood in the grislier way, (14) has only two meanings, and paraphrase requires appropriate tense/aspect on the embedded verb: we saw the spy while we walk towards the store; and we saw the spy who is walking towards the store.
When understood as a meaningful complex expression, as opposed to a mere string of words, (15) implies that the store is such that we saw the spy walk towards it. By contrast, (16) is as ambiguous as (14). But (17) is unambiguous in the way that (15) is.

(16) this is the store such that we saw the spy walking towards it
(17) this is the store such that we saw the spy walk towards it

Likewise, (18) can be only be understood as the question corresponding to (15):

(18) what did we see the spy walking towards

which is such that we saw the spy walk towards it? One can ask which thing is such that we saw the spy while walking towards it, or such that we that we saw the spy who was walking towards it. But these are different questions. As Chomsky notes, this suggests a common constraint on the meanings of relative clauses and questions.

When a child acquires English, she acquires a human language that connects meanings with pronunciations with in a certain way. This tells us something about what meanings are; human languages connect them with pronunciations with in limited ways. For example, even bracketing the lexical homophony of ‘saw’ and focusing on the past tense of ‘see’, the pronunciation of (14) has more meanings than the pronunciation of (15). Likewise, the pronunciation of (3) has two meanings,

(3) the guest is ready to please

which correspond to the meanings of the unambiguous (1) and (2).

(1) the guest is easy to please
(2) the guest is eager to please

In proposing that meanings are instructions for how to build mental symbols of a special sort, I am saying that human languages connect such instructions with pronunciations in these limited ways. An alternative hypothesis, discussed below, is that meanings are mappings from contexts to contents of some kind—often characterized in terms of Fregean senses, Russellian propositions, Tarskian satisfaction conditions, or some notion of possible worlds—and that human languages connect such mappings with pronunciations in constrained ways. But more can be said about what meanings are, in relatively neutral way, before turning to particular proposals.

### 1.2 Lexical Meanings and Combination: Arbitrary within Limits

Children acquire languages that permit two types of homophony, lexical and combinatorial. These subspecies of homophony are subject to different constraints. This provides another way of anchoring talk of meanings. As we’ll see, it also provides some initial motivation for describing meanings as composable instructions.

Given any overt lexical item L of a human language H, H connects the pronunciation π of L with some number of meanings. So initially, it might seem that lexical and combinatorial homophony are on a par: each overt pronunciation has n meanings for some n. But lexical homophony reflects the arbitrary character of lexical pronunciation-meaning pairs. The pronunciation of ‘cow’ could have gone with the meaning of ‘bare’, which could and did (also) go with the pronunciation of ‘bear’, which could have gone with the meaning of ‘chase’, etc. So if a given lexical pronunciation has two but not three meanings, this is unlikely to be a consequence of any interesting constraint.

Moreover, just as a particular speaker of English might not acquire the word ‘quasar’, a particular speaker might not acquire one of the meanings that other speakers connect with ‘bull’ or ‘sanction’. (Recall that as I am counting lexical items, each one connects a single meaning μ with a single pronunciation π, thereby connecting one way of understanding π with one way of pronouncing μ.) Even if there is a communal lexicon for English, a person can be a speaker of
English without encountering—much less creating an entry in memory for—each item on that long international list. So if a given speaker has a lexicon that connects a certain pronunciation with two but not three meanings, this fact would seem to be accidental in more than one way.

That said, there is another sense in which lexical pronunciation-meaning pairs are not arbitrary. It isn’t a matter of convention that children exposed to ordinary uses of ‘cow’ do not acquire a lexical item with the pronunciation *kaʊ* and the meaning of ‘mammal’, ‘set of hooves’, or ‘cow before I become a teenager and goat thereafter’; cp. Goodman (1954). The concepts that are naturally available to children presumably play a significant role in lexical acquisition; see, e.g., Bloom (200x). I’ll come back to this point, and to some grammatical constraints on lexical meanings. But in any case, we shouldn’t be surprised if limits on lexical meanings have their own character and etiology. For within the space of potential lexical meanings for a typical child, it does seem to be arbitrary—and perhaps conventional in Lewis’ (196x) sense—which ones get pronounced which ways in a particular community; see section two below.

By contrast, both combinatorial homophony and the ways it is constrained seem to reflect the nature of human languages. We could add another meaning to the pronunciation of ‘store’; and we could eventually subtract a meaning from the pronunciation of ‘saw’. But given a lexicon, we cannot make (14) less ambiguous or (15) more ambiguous.

(14) we saw the spy walking towards the store
(15) this is the store that we saw the spy walking towards

To take a simpler illustration, the meaning of (19) is intuitively conjunctive, in the sense of being paraphrasable with (19a) as opposed to other logically coherent alternatives;

(19) this is a cow we saw yesterday
(19a) This is both a cow and a thing we saw yesterday.
(19b) #This is either a cow or a thing we saw yesterday.
(19c) #This is a cow if we saw it yesterday.
(19d) #This is a thing that was a cow when we saw it yesterday.

Similarly, ‘brown cow’ is roughly synonymous with ‘cow and a brown one’. In later chapters, I argue that the meanings of ‘brown cow’ and ‘cow we saw yesterday’ are conjunctive in a way that is characteristic of phrasal meaning in general. But however we describe the significance of combining a word like ‘cow’ with an expression like ‘(that) we saw yesterday’, this mode of combination seems to be univocal, and not itself a source of ambiguity. For example, the homophony of ‘angler (who/that is) ready to eat’ is due to the homophony of ‘angler’ and ‘ready to eat’, not how ‘angler’ combines with the relative clause. And while children could easily acquire a variant of English in which the pronunciation of ‘ready’ has an extra meaning, or the pronunciation of ‘carp’ has only the meaning of ‘trout’, they do not naturally acquire variants in which combining a word with a relative clause has some non-actual significance.

We can invent a pseudo-English in which (19) has the meaning of (19c) or (19d) or both, once we already have a human language. But it isn’t an accident of English that (19) has only the meaning indicated with (19a). Even if the significance of combining a word with a relative clause is not fixed by the innate endowment that children use to acquire languages, the non-ambiguity of such combination is not arbitrary in the way that absence of lexical homophony is. There is a sense in which it is an accident of English that no meanings of ‘bucket’, ‘stamp’, and ‘jump’ share a pronunciation; cp. ‘seau’, ‘sceau’ and ‘saut’ in French. But it is no accident that (19) is not a homophone that also has the meaning of (19b).

Acquiring a language that allows for certain homophones but not others is an impressive accomplishment—especially for an immature primate—even if one can safely assume that
modes of grammatical combination have univocal significance. Put another way, if children could acquire languages in which ways of combining lexical items are ambiguous, then it would be even harder to explain how children do acquire languages in which combinatorial homophony is limited in the ways it is; see Higginbotham (1985). Indeed, given the arbitrary character of atomic meaning-pronunciation pairs, acquiring a lexicon is hard enough even if one assumes that (i) lexical items can be combined in a small number of ways, each corresponding to a specific kind of significance, and (ii) combinatorial homophony is due to possibility of combining lexical items in different ways.

If human languages conform to these assumptions, say because children select languages from a space of options partly defined by (i) and (ii), then each way of combining lexical items has only one meaning. Two or more ways of combining the same lexical items—i.e., two or more complex expressions—may share the same pronunciation, because “linearizing” the structured expressions may yield the same string in each case. Correlatively, a single string of words may be classified as the linearization of two or more structure expressions. In terms of a familiar kind of example, (20) can be understood in two ways, indicated with (20a) and (20b).

(20) a spy saw a man with a telescope
(20a) A spy saw a man by using a telescope.
(20b) A spy saw a man who had a telescope.

These two meanings pretty clearly reflect two ways of structuring the words in (20): one way in which ‘saw a man with a telescope’ has ‘saw a man’ and ‘with a telescope’ as constituents, as shown in (20c); and one way in which ‘man with binoculars’ is a constituent, as shown in (20d).

(20c) {{a spy} [[saw [a man]] [with [a telescope]]]]
(20d) {{a spy} [saw [a man [with [a telescope]]]]}

The curlyness of the outermost brackets, akin to the capitalization and periods in (20a) and (20b), is just a reminder that these structured expressions are understood as complete sentences.7

Theorists can hypothesize that a single way of combining lexical items can have more than one meaning. But this abandons the following attractive idea: when lexical items are combined in a particular way, the result is an expression whose meaning is determined by those lexical items and that mode of combination. Moreover, if a string of words can be ambiguous in a way not resolved by disambiguating lexical items and fixing on a particular way of combining them, then we still need an explanation for the phenomenon of constrained homophony.

String (20) cannot be understood as a sentence that can be paraphrased with (20c).

(20c) #A spy saw a man and had a telescope.

So if a combination of lexical items can have more than one meaning, then even given that the constituency structure of (20c) supports the meaning indicated with (20a), one wants to know why this constituency structure cannot also support the meaning indicated with (20c). As (20b) illustrates, ‘with a telescope’ can be understood along the lines of ‘(who) had a telescope’. This highlights the question of why ‘with a telescope’ cannot also be understood this way in (20c).

---

7 Compare the ‘s’ in [s [NP a spy] [VP saw [NP a [N man [PP with [NP a telescope]]]]]]. See Kayne (200x) on the notion of linearization. Chomsky (1957, 1964, 1965) spoke of sentences as “strings of formatives” that have certain “structural descriptions;” see Pietroski (2015) for discussion highlighting connections to Goodman (1954). The idea was that certain strings are classified as sentential by structural descriptions that included a privileged symbol ‘S’, intimating sentences but also the “Start” symbol in a system of rewrite rules like “S → NP (aux) VP”. Since current grammatical theories do not take the notion of sentence as basic, one has to say what sentences are. But as we’ll see, I do not posit any categorical distinction between sentences and phrases. On my view, sentences can be phrases that interface with the rest of cognition in a distinctive way, as opposed to expressions of a special linguistic type.
I’ll soon take up the question of why \((20\alpha)\) does not have the meaning of \((20c)\) and no other. Here, I just want to stress the cost saying that a single way of combining lexical items can have more than one meaning. Whatever one says about the constituency structures that get assigned to strings like \((1), (2), (6), (15),\) and \((19),\)

\(1\) the guest is easy to please  
\(2\) the guest is eager to please  
\(6\) ducks that are easy to eat with a fork are eager to eat with a spoon  
\(15\) this is the store that we saw the spy walking towards  
\(19\) this is a cow we saw yesterday

each of the corresponding sentences fails to have the meanings that it fails to have. Interesting non-homophony is ubiquitous. So it seems reasonable to conclude—at least as an idealization to be relaxed only in limited ways—that each combination of lexical items has only one meaning.

This makes it possible to say a little more about what I take meanings to be. Each expression of a human language connects a meaning \(\mu\) with a pronunciation \(\pi\). Lexical items connect their meanings with pronunciations in ways that are arbitrary in familiar respects, but limited in others; and the idiosyncracies of lexical meaning-pronunciation connections allow for one kind of homophony. Complex expressions, which are results of combining lexical items in certain ways, connect their meanings with pronunciations in structure-dependent ways. This creates the possibility of combinatorial homophony, within limits imposed by lexical items and how they can be combined. Any account of linguistic meanings must accommodate these basic facts regarding the two kinds homophony that human languages exhibit. There are various ways of saying how lexical and phrasal meanings differ. But let me end this subsection by advertizing my own proposal, without yet arguing that it is better than alternatives.

Suppose that humans, and perhaps many animals, enjoy mental symbols that have the following characteristics: they can be used to think about things in certain ways; and while the symbols may be segregated in ways that limit how they can be combined, each of them can be combined with some others, so that a symbol used to think about things in one way (e.g., as cows) can be part of a symbol used to think about things in a distinct but analytically related way (e.g., as brown cows). Call such mental symbols concepts, leaving further details for chapter one. In my view, the meaning of a lexical item \(L\) is an instruction for how to access a concept from a mental address that is linked to the pronunciation of \(L\); where accessed concepts are of a special sort, and often distinctively human, but the mapping from accessed concepts to pronunciations is arbitrary. By contrast, I think the meaning of a phrase is an instruction for how to form a concept by executing sub-instructions and combining the results in a certain way; where the relevant mode of composition may not be distinctively human, but the mapping from ways of forming expressions to ways of combining concepts is fixed, leaving no room for arbitrary variation.

I take no stand on how such instructions are implemented. Though in a von Neumann machine, strings of numerals are used to access and perform operations on other strings. In this sense, ‘1101011010’ can be an instruction that is executed by performing operation six \((110)\) on the number stored in \((1)\) register twenty-six \((011010)\), while ‘0100110101’ is executed by performing operation two \((010)\) on the number \((0)\) fifty-three \((110101)\). A biologist can also describe strings of DNA—or corresponding RNA sequences, bounded by “start” and “stop” codons—as instructions for how to build proteins by attracting amino acids that can combine in certain ways; and such descriptions can be useful, even if they are ultimately inessential. Instructions for how to build concepts may be representational in way that strings of DNA are not. But my goal is to say what meanings are, not to reduce talk of representations.
The details regarding accessible concepts and combinatorial operations are for chapters one, five, and six. But I will argue that the operations, few in number, only generate *predicative* concepts that exhibit *conjunctive* structure. In my view, executing the meaning of ‘brown cow’ yields a concept with which a thinker can classify something as—or predicate of something that it is—both a cow and a brown one. Executing the meaning of ‘(that) we saw yesterday’ yields a correspondingly conjunctive concept with which a thinker can classify one or more things as things we saw yesterday. However, it takes work to describe (i) the kinds of concepts that ‘see’, ‘yesterday’, ‘cow’, and ‘brown’ can access, and (ii) and how such concepts can be combined.

### 1.3 Phrasal Meanings and Constituency: Composition and Specification

Details aside, the phenomenon of constrained homophony suggests that when lexical items are combined in a particular way, the result is an expression whose sole meaning is determined by those lexical items and that mode of combination. But there is little agreement on the character of this determination, and there are many definable notions of “compositionality”.

There can also be terminological disputes between those who focus on human languages and those who focus on the general conditions that *any* language must satisfy if it generates endlessly many interpretation-signal pairs of some kind. But let me flag an important dimension of disagreement about meanings before returning to the task of characterizing them in a suitably neutral way.

I think human languages respect a demanding constraint according to which the meaning of a complex expression CE has the meanings of its constituents as *parts*. More precisely: if CE is formed from the simpler expressions E₁…Eₙ in a certain way W, then CE has a meaning of the form \( \mu(W)[\mu(E₁)…\mu(Eₙ)] \); where \( \mu(\alpha) \) is the meaning of \( \alpha \), and \( \mu(W) \)—the meaning of a certain mode of grammatical combination—builds complex meanings from simpler ones. In particular, if CE is a phrase formed by combining E₁ and E₂ in a way that signifies a certain kind of conjunction, then \( \mu(CE) \) can be the tripartite instruction \( \text{JOIN}[\mu(E₁), \mu(E₂)] \); where this meaning is executed by applying a certain conjunctive operation to the results of executing the constituent meanings/instructions. But the idea that expression meanings exhibit part-whole relations—mirrored by the constituency structures of expressions—is controversial.

There is more agreement on the weaker requirement that \( \mu(CE) \) be *specifiable* as the value of some operation, determined by the relevant mode of grammatical combination, given the meanings/values of the constituents of CE as operands; see Davidson (1967b), Lewis (1972), Montague (1974). This does not imply that \( \mu(CE) \) has parts, much less parts that are meanings of parts of CE. If CE is formed by combining E₁ and E₂ in way W, the weak requirement is compatible with \( \mu(CE) \) being an object—e.g., a truth value or a set—that has no parts, or at least no parts that include \( \mu(E₁) \) and \( \mu(E₂) \). As discussed in chapter two, \( \mu(W) \) might be the operation of function-application, while \( \mu(E₁) \) is a function from entities in a domain that includes \( \mu(E₂) \) to truth values. And one can posit further constraints, on mappings from expressions to specifiable “semantic values,” without saying that meanings have parts. I doubt that ordinary expressions (as opposed to idealized concepts) bear any interesting relation to the posited semantic values, much less a relation not mediated by composable meanings; see chapters three and four. But whether or not phrasal meanings have forms that specifications should reflect, some specifications are better than others. And this is enough to complete our initial characterization of meanings.

---

8 For discussion, see Szabo (2000), Pelletier (199x), Pagin (200x), Hodges (19xx), Jacobson and Barker (201x).

9 In this sense (and others) my conception of meaning is akin to Frege’s (1892) notion of Sinn; see also Church (1951), Hory (2007), and the notion of “procedure” discussed in section two. Though one can posit structured meanings that are not *Sinnen*; see, e.g., Russell (19xx), Cresswell (1985), Soames (1987), and King (2007).
Recall that (20) is roughly synonymous with (20a) and (20b),

(20) a spy saw a man with a telescope
   (20a) A spy saw a man by using a telescope.
   (20b) A spy saw a man who had a telescope.
   (20c) #A spy saw a man and had a telescope.

which correspond to sentences (20α) and (20β), neither of which can be paraphrased with (20c),

(20α) {[a spy] [saw [a man]] [with [a telescope]]}
(20β) {[a spy] [saw [a man [with [a telescope]]]]}

even given loose standards of paraphrase. In this sense, we can use (20a-c) to indicate sentential meanings. Indeed, we can use (20a) to specify the meaning of the unambiguous string (21).

(21) a spy saw a man by using a telescope
(21a) A spy saw a man by using a telescope.

Similarly, (22) is boringly correct, modulo concerns about the context sensitivity of tense.

(22) The sentence ‘A spy saw a man by using a telescope.’
means that a spy saw a man by using a telescope.

But there is a perfectly good sense in which (20a) specifies the meaning of (20α) better than (20α) specifies its own meaning, even if (20a) and (20c) are not strictly synonymous.

There are at least two dimensions along which proposed specifications of meanings can be ranked. One corresponds to the vague, context sensitive, and gradable notion of good paraphrase (translation, or Davidson’s [1968] notion of “samesaying”). Another corresponds to the vague, context sensitive, and gradable notion of theoretical illumination. Saying what makes one specification better than another is tantamount to saying what meanings are. So we shouldn’t expect theory-neutral ground rules for choosing between alternative specifications. But at least prima facie, some specifications are better/worse than others.

We needn’t think that (23) and (24) are perfectly synonymous to think that each paraphrases the other pretty well, and that (25) is not even a minimally decent paraphrase of (23).

(23) Georgiana is eager to please.
(24) Georgiana is eager to be one who pleases others.
(25) Georgiana is eager to be one whom others please.

It seems that (23) and (24) paraphrase, or at least imply each other. It is hard to describe a situation in which exactly one of these sentences could be used to describe Georgiana correctly. By contrast, (23) and (25) have very different implications. I return to the notion of implication more than once; though for present purposes, any pre-theoretic conception will suffice.

The phrases ‘man who had a telescope’ and ‘man with a telescope’ are not synonymous, given the possibility of men next to telescopes owned by others. But this difference seems subtle compared with the sharp contrast between (20b) and (20c). In this sense, (20b) is better than (20c) as a specification of what (20β) means. When it comes to hard cases, such judgments may reflect a theoretically intractable mix of considerations; see, e.g., Burge (197x). But we don’t need a general theory of paraphrase to use (20a) and (20b) as specifications of the meanings of (20). And we don’t need a theory of what makes one specification more illuminating than another to see that in this respect, (24) is better than (23) as a specification of what (23) means. Some illumination is better than none. To be sure, (23) paraphrases itself perfectly, as does (26).

(26) Georgiana is easy to please.

But if one wants to know how (23) and (26) differ in meaning, and why ‘ready to please’ is ambiguous, then using (23) and (26) to specify their own meanings provides no illumination.
Returning to (20), we can stipulate that μ(\{a spy\} \{saw [a man] with [a telescope]\}) is the meaning of (20α). But there is a perfectly good sense in which (20a) specifies this meaning better than (20α) does, and (20b) specifies the meaning of (20β) better than (20β) does. We can use the ordinary sentences to ask why (20α) has the meaning of (20α) as opposed to (20c).

To sharpen this question, consider the following “regimentations” of (20b) and (20c).

(20b) \(\exists x\{\text{Spy}(x) \land \exists y\{\text{Saw}(x, y) \land \text{Man}(y) \land \exists z\{\text{Had}(y, z) \land \text{Telescope}(z)\}\}\}\)

(20c) \#\exists x\{\text{Spy}(x) \land \exists y\{\text{Saw}(x, y) \land \text{Man}(y)\} \land \exists z\{\text{Had}(x, z) \land \text{Telescope}(z)\}\}

By stipulation, these sentences of an invented formal language have truth conditions—satisfaction conditions that determine truth values—that can be recursively specified in the usual way.\(^\text{10}\) We can be agnostic, or agree to disagree, about how (20b’) and (20c’) are related to (20b) and (20c). The “formal translations” do not have meanings if meanings are composable instructions; and as discussed in chapters three and four, one can doubt that sentences of human language have truth conditions. But the invented sentences, which specify truth conditions in ways that make certain logical relations explicit, can be used to model certain aspects of what human sentences do and do not mean. Even if (20b’) does not have the meaning of (20b), (20b’) can be used to describe this meaning in a nontrivial way.

I can view (20b’) as an initial suggestion about the form of the meaning that (20b) connects with a pronunciation, or perhaps the form of an idealized concept that could be built by executing this meaning. Others can say that the meaning of (20b) has the “semantic value” that (20b’) specifies. Either way, (20b’) can be used to describe the meaning of (20β) in a way that highlights a respect in which ‘with a telescope’ is like ‘\(\exists z\{\text{Had}(x, z) \land \text{Telescope}(z)\}\)’.

(20β) \{[a spy] \{saw [a man [with [a telescope]]]\}\}

In (20β), the prepositional phrase is understood as a predicate that classifies things according to whether or not they had a telescope. So one wants to know why this phrase is not understood the same way in (20α), yielding the meaning specified with (20c/20c’).

(20α) \{[a spy] [[saw [a man]] [with [a telescope]]]\}

We can imagine languages in which ‘saw a man’ is like ‘\(\exists y\{\text{Saw}(x, y) \land \text{Man}(y)\}\)’ in being understood as a predicate that classifies things according to whether or not they saw a man, and combining phrases is like adjoining invented formulae, with the result that (27)

(27) \{[saw [a man]] [with [a telescope]]\}

is understood as a predicate that classifies things according to whether or not they saw a man and/or a telescope. English is not such a language. But if (27) is not like (28), what is it like?

(28) \(\exists y\{\text{Saw}(x, y) \land \text{Man}(y)\} \land \exists z\{\text{Had}(x, z) \land \text{Telescope}(z)\}\)

This question can also be sharpened. Sentence (29) might be regimented with (20c’),

(29) A spy with a telescope saw a man.

ignoring order of conjuncts. So perhaps ‘with’ has distinct “possessing” and “by-using” construals. But even so, why does (20α) have the meaning of (20b) rather than (29a)?

(29a) A spy with possessing a telescope saw a man.

(29b) A spy with by-using a telescope saw a man.

---

\(^\text{10}\) See chapter two. Here, let’s not worry about vagueness or context sensitivity. In particular, the tense of ‘had’ is not the issue; ‘man with a dog’ can be understood as ‘man who has a dog’. We can also make (20b’) and (20c’) look more like instances of ‘μ(W)[μ(SE₁), μ(SE₂)]’ by rewriting with dyadic/restrictable quantifiers:

\(\exists x\{\text{Spy}(x), \exists y\{\text{Man}(y)\}, \exists z\{\text{Telescope}(z), \text{Had}(y, z)\}, \text{Saw}(x, y)\}\}:

\(\exists x\{\text{Spy}(x), \exists y\{\text{Man}(y), \text{Saw}(x, y)\}, \exists z\{\text{Telescope}(z), \text{Had}(x, z)\}\}\).
We can say, correctly enough, that the English verb phrase (27) has the meaning of (29b). But one wants to know how the lexical items and grammatical structure of (27) determine this meaning. Likewise, specifying the meaning of (27) as \( \mu([\text{saw a man}] \text{ [with a telescope]]) \) leaves it mysterious how this meaning—and not, say, the meaning of (30)—

(30) saw a man and had a telescope

is determined by the lexical items and grammatical structure of (27).\(^{11} \)

Luckily, there are better ways of specifying \( \mu([\text{saw a man}] \text{ [with a telescope]]) \). The meanings of (20) can also be described with (20aa) and (20bb),

(20) a spy saw a man with a telescope
(20a) A spy saw a man by using a telescope.
(20aa) Some event was a spy seeing a man and done with a telescope.
(20b) A spy saw a man who had a telescope.
(20bb) Some event was a spy seeing a man who had a telescope.

which paraphrase (20a) and (20b) in prolix fashion; see Davidson (1967a). In most contexts, paraphrasing (31) with (31a) would be odd. Indeed, an editor might well replace (31a) with (31).

(31) A spy saw a man.
(31a) Some event was a spy seeing a man.

Needless prolixity distacts. But in some contexts, (31a) is better than (31) as a specification of what (31) means, and (20aa) is better than (20a) as a specification of what (20a) means.

(20aa') \( \exists x \exists y \{ \text{Spy}(x) \text{ & } \exists y[\text{See}(e, x, y) \& \text{Past}(e) \& \text{Man}(y) \text{ & } \exists z[\text{With}(e, z) \& \text{Telescope}(z)] \} \)

where ‘\( \text{See}(e, x, y) \)’ is true of ordered triples \( <e, \alpha, \beta> \) such that \( e \) is an event of \( \alpha \) seeing \( \beta \). From this perspective, ‘saw a man’ is unlike ‘\( \exists y[\text{Saw}(x, y) \& \text{Man}(y)] \)’ and more like ‘\( \exists y[\text{See}(e, x, y) \& \text{Past}(e) \& \text{Man}(y)] \)’, which is true of ordered pairs \( <e, \alpha> \) such that \( e \) is an event of \( \alpha \) seeing a man. The meaning of (20b) can be specified in similar fashion.

(20bb') \( \exists x \exists y \{ \text{Spy}(x) \text{ & } \exists y[\text{See}(e, x, y) \& \text{Past}(e) \& \text{Man}(y) \text{ & } \exists z[\text{With}(y, z) \& \text{Telescope}(z)] \} \)

That is a step in a good direction. As discussed below, it also lets us start to explain the sense in which (20a) implies (31). But if ‘saw a man’ is like a formula with two unbound variables, one of which corresponds to the ‘experiencers’ in events of seeing, then we still face the question of why ‘saw a man with a telescope’ cannot be understood along the lines of ‘\( \exists y[\text{See}(e, x, y) \& \text{Past}(e) \& \text{Man}(y) \text{ & } \exists z[\text{With}(x, z) \& \text{Telescope}(z)] \)’. If the prepositional phrase can be associated with the ‘e’-position as in (20aa’) or the ‘y’-position as in (20bb’), then why can’t it be associated with the ‘x’-position as in (20cc’)?

(20cc') \( \exists x \exists y \{ \text{Spy}(x) \text{ & } \exists y[\text{See}(e, x, y) \& \text{Past}(e) \& \text{Man}(y) \text{ & } \exists z[\text{With}(x, z) \& \text{Telescope}(z)] \} \)

So perhaps (20aa’’) is an even better way of specifying the meaning of (20a);

---

\(^{11}\) *Pace* Horwich (1997, 1998), it’s not enough to display (20a) while saying that its elements and modes of combination mean what they do; see Pietroski (2000b, 2005b) for extended discussion.
(20aa”) \( \exists \{ \exists x [ \text{Experiencer}(e, x) \& \text{Spy}(x)] \& \\
\exists y [ \text{Pssee}(e, y) \& \text{Past}(e) \& \text{Man}(y) \& \exists z [ \text{With}(e, z) \& \text{Telescope}(z)]] \}
\]

where ‘Pssee(e, y), ‘P’ connoting ‘Passive’, is true of pairs \( \langle e, \beta \rangle \) such that \( \epsilon \) is an event of seeing \( \beta \). On this view, ‘saw a man’ is more like \( \exists y [ \text{Pssee}(e, y) \& \text{Past}(e) \& \text{Man}(y)] \)’, whose only variable position corresponds to events as opposed to experiencers; cp. Kratzer (1996).

I return to further details and emendations—including the option of replacing ‘Pssee(e, y)’ with ‘Pssee(e) & Theme(e, y)’—in chapters one, five, and six. But to advertise, I think (20aa”) illustrates a general pattern: lexical meanings have at most two variable positions; phrasal meanings are conjunctive, and they have only one variable position. This raises further questions. If the lexical item ‘see’ accesses a concept that is like ‘Pssee(e, y)’ or ‘Seeing(e)’ in having an event position, how is such a concept acquired? Do infants naturally employ such concepts and assign them to lexical addresses; or do accessed concepts have to be introduced, perhaps by using concepts that are more like ‘See(x, y)’ or ‘See(e, x, y)’? What, if anything, are the accessed concepts true of? Addressing these questions can make certain proposals about the meaning of ‘see’ more or less plausible. Indeed, that is the point. On any reasonable account of what makes one meaning specification of better than another, ‘See(x, y)’ seems inferior to ‘See(e, x, y)’, which may be inferior to an option like ‘Pssee(e, y)’ or ‘Seeing(e)’; though evaluating options leads to further questions, whose answers have implications that we cannot foresee.

Whatever lexical meanings are, they conspire with grammatical structure to determine phrasal meanings in ways that are not obvious. Even if there turns out to be a viable alternative to specifying verb meanings in terms of events, that would only reinforce the general point: specifying \( \mu \) (‘see’) in an illuminating way takes work, and likewise for specifying the significance of combining ‘see’ with other lexical items to form complex expressions like [[saw [a man]] [with [a telescope]]]. Examples like (14) introduce further twists,

(14) we saw the spy walking towards the store
   (14a) We saw the spy while we were walking towards the store.
   (14b) We saw the spy who was walking towards the store.
   (14c) We saw the spy walk towards the store.

raising questions about how to best specify \( \mu \) (‘walk’). We have to discover the best ways of describing meanings, just as we have to discover the best ways of describing how lexical items can be combined (and how planets orbit their suns); cp. Chomsky (1965).

If this point is accepted, then I don’t think we need to start with more controversial claims about the meanings of human linguistic expressions. Once we appreciate the phenomenon of constrained homophony, ordinary linguistic competence provides a good initial sense of what meanings are. This initial sense can be boosted by using invented languages to sharpen questions about why certain arrangements of lexical items have the meanings they do, and how meanings should be specified for purposes of addressing such questions. In this way, we can start with a relatively uncontroversial conception of linguistic meanings that is arguably richer than our presuppositional conception of butterflies. Then we can let the chips fall where they may with regard to how meaning is related to concepts, truth, and communication.

2. It’s All About I-languages

There is, however, a sense in which the topic may still be unclear. I have assumed that human languages generate expressions, each of which connects a meaning \( \mu \) with a pronunciation \( \pi \); where each expression is either a lexical item, which connects its meaning with a pronunciation in a way that is largely arbitrary, or a structured combination of lexical items. But I didn’t say much about human what languages are, apart from an introductory suggestion that they are
biologically implemented procedures. That suggestion needs to be spelled out, especially since I also spoke of languages—some merely possible, some already invented—that are like human languages in certain respects yet crucially different in other respects. Formal languages are often described as sets of expressions, and there is a well-known tradition of taking human languages to be special cases of sets; see, e.g., Lewis (1975). In my view, human languages are I-languages in Chomsky’s (1986) sense. But this technical locution has come to be used in many ways. So let me explain how it is used here, and then say why I take human languages to be biologically implemented I-languages, as opposed to sets of expressions.

2.1 ‘I’ before ‘E’ especially after ‘Ch’

Often, we want to describe human languages as special cases of languages in a broader sense. We can and do speak of mathematical languages, bee languages, languages of thought, etc. So at least initially, let’s adopt a generous conception of languages that covers anything that somehow connects interpretations of some kind with signals of some kind. Human languages can be described as special cases that connect interpretations of a particular sort (meanings) with signals of a particular sort (pronunciations). This leaves room for many proposals about the respects in which human languages are distinctive. But languages need not share a common nature. Similarly, various things fly: birds, bats, butterflies, flies, propeller planes, jumbo jets, helicopters, etc. One can debate whether high-tech sailboats briefly fly, or whether gliding with a jet pack counts. But the uncontroversial cases already include things of disparate sorts. Whatever it is to be capable of flight, this capability is “realized” in different ways by things of different kinds. The things that connect interpretations with signals are presumably at least as diverse.12

We can, however, say that all languages have expressions. In simple cases, expressions can be identified with interpretation-signal pairs; though an expression can also be a “syntactic structure” that is connected to both a signal and an interpretation. An expression might pair a certain sound or gesture or inscription with a certain object or property or concept. Or the expressions of a language might connect instructions of a specific kind with strings of letters; where these instructions are themselves expressions of another language that connects electrical impulses with strings of ‘1’ s and ‘0’ s in a certain way. This leaves it open what it is for a language to “have” expressions. In particular, some languages are procedures that generate their expressions, while other languages are sets that contain their expressions as elements.

Given this generous conception of what a language is, the space of possible languages can be divided in many cross-cutting ways. We can classify languages according to how many expressions they have (e.g., finitely or unboundedly many). We can also distinguish explicitly invented languages from their natural counterparts, and then distinguish the latter in terms of the mechanisms required to detect the relevant signals, which need not be acoustic or within the range of normal human perception. Or we might distinguish abstract procedures that generate interpretation-signal pairs from all other languages, including any sets that are the extensions of such procedures, and then classify the former in ways familiar from the study of computation.

---

12 I don’t deny that speakers can use ‘thing that flies’ to express a concept that applies to birds but not airplanes. We have some ungenerous concepts. Similarly, we can use ‘water’ to express a concept that applies to samples of H2O but not stuff of the sort that Putnam (1975) imagined. But I don’t see any good reasons for thinking that ‘water’ is a “natural kind term,” with an extension that excludes any watery stuff that isn’t H2O (modulo impurities), while ‘flies’ does not likewise exclude airplanes; see Chomsky (2000b), Pietroski (2005b). My own intuitions, as a native speaker of English, suggest that ‘water’ is not limited in this severe way; though in a conversation where natural kinds are what matter, I certainly feel the inclination to use ‘water’ (and ‘flies’) in the restricted way.
Quite apart from the study of human languages, it is clear that not all languages are sets. When we invent a language that has infinitely many expressions, we do so by using finitely many expressions (of an already available language) to specify some procedure \( P \) that generates infinitely many expressions; where ‘generates’ is used in an abstract sense, familiar from mathematics, without implying episodes of generating the expressions. If other procedures generate the same expressions, but it doesn’t matter which procedure one uses to specify those expressions, it can be useful to define a certain language \( L \) as the set of expressions generated by \( P \)—and hence by any procedure that is “extensionally equivalent” to \( P \). But it hardly follows that \( P \) is not a language. Whatever interpretations \( L \) connects with signals, \( P \) connects those interpretations with those signals. Indeed, \( L \) is specified as \( \text{Extension}(P) \).

The inventors of modern logic were clear about this. Frege’s (1892a) notion of a Function was procedural. He had the notion of a set of ordered pairs with no elements \( <\alpha, \beta> \) and \( <\alpha, \gamma> \) such that \( \beta \) and \( \gamma \) differ. But according to Frege, each Function has a “course of values,” or what we would call an extension. If two Functions share a course of values, we may want to ignore ways in which the procedures differ and focus on what they have in common; Frege (1892b) introduces his notions of sense (\( \text{Sinn} \)) and semantic value (\( \text{Bedeutung} \)) in a related way. But for Frege, the basic logical objects are Functions that determine courses of values. And this idea can be isolated from his more problematic talk of Functions as “unsaturated;” see chapter two. Church (1941) distinguished “functions in intension” from “functions in extension.” This distinction, introduced in the context of discussing computability in Turing’s (193x) sense, is also the source of Chomsky’s (1986) contrast between I-languages and E-language. So it is worth noting that both Church and Chomsky used ‘intension’ as a variant of ‘procedure’.

Consider the set of ordered pairs \( <x, y> \) such that \( x \) is a whole number, and \( y \) is the absolute value of \( x - 1 \). This infinite set, \{ …, (−2, 3), (−1, 2), (0, 1), (1, 0), (2, 1), … \}, can be described in many ways. Using the notion of absolute value, one can say that \( F(x) = |x - 1| \). Or one can use the notion of a positive square root: \( F(x) = \sqrt{x^2 - 2x + 1} \). These descriptions correspond to different procedures for computing a value given an argument; and a mind might be able to execute one algorithm but not the other. Church put the point in terms of his lambda expressions, discussed in chapter two: one can interpret such expressions intensionally, saying that \( \lambda x. |x - 1| \) and \( \lambda x. \sqrt{x^2 - 2x + 1} \) are distinct but extensionally equivalent procedures; or one can interpret lambda expressions extensionally, saying that \( \lambda x. |x - 1| \) is the same set as \( \lambda x. \sqrt{x^2 - 2x + 1} \). But to specify the space of computable functions, one needs the first interpretation. And the procedural construal lets us say that \( \text{Extension}[\lambda x. |x - 1|] = \text{Extension}[\lambda x. \sqrt{x^2 - 2x + 1}] \). So the set-theoretic construal is inessential, though sometimes convenient.\(^{13}\)

---

13 The key proof in Church (1941) employs what he calls “the calculus of \( \lambda\delta \)-conversion.” He invents this formal language in stages. But at the outset, he introduces the intensional/extensional contrast and says where he is heading. In the calculus of L-conversion and the calculus of restricted \( \lambda\delta \)-conversion, as developed below, it is possible, if desired, to interpret the expressions of the calculus as denoting functions in extension. However, in the calculus of \( \lambda\delta \)-conversion, where the notion of identity of functions is introduced into the system by the symbol \( \delta \), it is necessary, in order to preserve the finitary character of the transformation rules, to formulate these rules that an interpretation by functions in extension becomes impossible. The expressions which appear in the calculus of \( \lambda\delta \)-conversion are interpretable as denoting functions in intension of an appropriate kind. (p.3, my italics).

Crucially, a function in intension is not a set of possible worlds. We can use such sets to model certain equivalence classes of procedures; cp. Church (1951). But distinct procedures can have the same extension at every possible world. Though as Church stressed, there are various ways of individuating procedures (which he associated with meanings), and hence many notions of functions in intension “involving various degrees of intensionality.”
Chomsky (1986) describes an I-language as a procedure that connects interpretations with signals; an E-language is a language in any other sense—a set of interpretation-signal pairs, a cluster of dispositions to act in certain ways with regard to certain utterances, or whatever else a language might be taken to be. As Chomsky notes, it is obvious that there are many I-languages: to even specify a set with endlessly many elements, one must somehow specify a procedure that determines the set. This point applies to children as well as theorists. And as illustrated in section one, acquiring English seems to be a matter of acquiring a procedure that connects meanings with pronunciations in a particular way. Barring especially liberal procedures—e.g., connect any μ with any π—a procedure that generates μ-π pairs in a certain way will thereby generate some such pairs and not others. A procedure can also be acquired, in roughly the sense that programs can be downloaded, even if we don’t know how human biology supports this process. So I take it to be obvious that children acquire I-languages.

As Chomsky notes, the I-languages that children acquire are plausibly individualistic (as opposed to “public languages” that a speaker may never fully acquire) and internalistic—at least in the sense that molecular duplicates in different environments would have the same I-language, even if they used it to talk about different things. But while ‘I-’ connotes these further hypotheses about the expression-generating procedures that children acquire, these additional claims can be challenged; they are not stipulations about what can count as an I-language. Nor is there any stipulation that children do not acquire E-languages. Given a generous notion of acquisition, one can say that as children acquire I-languages, they acquire certain E-languages.

For example, a disposition to connect English pronunciations (up to some length) with their meanings can be an E-language, and a child who acquires English acquires such a disposition. If there is evidence that speakers of English connect the same meanings with the same pronunciations via procedures that differ in kind, or if we meet aliens who use a massive phrase book to communicate with us, it may be useful to speak of E-English. Though it’s hard to see any explanatory role for such an E-language as opposed to generative procedures that happen to be equivalent in certain respects. When children acquire English, they acquire I-languages that may differ in various ways. But on the whole, these procedures connect meanings with pronunciations in remarkably similar ways.

One can hypothesize that children acquire expression-generating procedures that determine sets of meaning-pronunciation pairs, and that reference to these sets/E-languages will play a role in our best account of what meanings are. But in my view, both parts of this hypothesis are false. In §2.2, I address the second part, in connection with Lewis’ (197x) claims about how human languages are related to conventions of “truthfulness and trust.” Here, let me stress that one can posit I-languages without supposing that they determine sets of expressions.

Talk of procedures is often introduced, as it was above, with examples of arithmetic operations that map natural numbers to natural numbers. Given this well-defined domain/range, the corresponding procedures have extensions, as Church noted. But the elements of these extensions are as abstract as the computational procedures. So once we start talking about physical implementations—e.g., detecting a certain shape in a given section of a tape, and responding in a certain way—we need to idealize away from various performance limitations.  

---

14 Ludlow (2011) introduces the notion ‘Ψ-language’ to explicitly cancel the further connotations (and invite an externalistic alternative). But in my view, his Ψ-languages just are I-languages by a slightly different name.

15 A calculator may impose an upper bound on which numbers can be represented, yet still implement procedures that are unbounded and constrained in ways that lookup tables are not. Kripke (1983) makes a certain form of rule-following skepticism vivid. But idealization remains legitimate; see Pietroski and Rey (1995) for discussion.
More importantly, with respect to the study of human languages, we don’t start out with axioms that tell us how to individuate meanings and pronunciations. And examples like (32)

(32) *the child seems sleeping

(32a) The child seems to be sleeping.
(32b) #The child seems sleepy.

tell against any simplistic analogies to invented languages in which a string of atomic symbols has an interpretation if and only if the string is “well-formed” in some independently specifiable sense; see Chomsky (1965), Higginbotham (1985).16

The asterisk indicates the unacceptability of (32). But this string is still comprehensible. Indeed, competent speakers of English know that it is unambiguous, with the meaning of (32a) as opposed to (32b). In acquiring English, one acquires a capacity to understand (32) in a certain way, and not mere a disposition to “repair” the string; compare the ill-formed (33),

(33) ∃x[Fx & ∀y[Rxy ⊃ ∃z[Szx]]

which can repaired by adding a right bracket, at the end or after the second occurrence of ‘y’. For whatever reason, most Americans—with western Pennsylvania and eastern Ohio providing contrast—find it hard to hear (34) as meaning that the lawn needs to be mowed.

(34) *the lawn needs mowed

But even for those of us who hear (34) as defective, it is not word salad in the way that (7) is;

(7) *we been have might there

see Lasnik (200x) for discussion. So even setting aside the dialectical variation that makes appeal to a single set of English expressions seem silly, it isn’t clear that any speaker of English has acquired an I-language that determines a particular set of meaning-pronunciation pairs.

Let Σ be the pair consisting of the meaning of (32a) and the pronunciation of (32). Each set either has Σ as an element, or it doesn’t. But this seems like a poor way to think about how speakers of English are related to Σ. The question of whether Σ is “in” my native language seems ill posed. It’s not as if Σ is either a variant pronunciation of (32a) or a foreign string that I can somehow decipher. One might reply that each I-language determines a set of triples <μ, π, n>; where n is a number, 0 < n ≤ 1, that reflects a degree of grammaticality. But while theoretically interesting notions of grammaticality may be gradable, my understanding of (32) doesn’t seem to be. Linguists may want to model grades of grammaticality, which may reflect aspects of how expressions interact with various systems used in the comprehension/production of speech. But children don’t need procedures that generate <μ, π, n> triples. Acquiring a procedure that

---

16 Chomsky (1957), originally lecture notes for an undergraduate course, adopts a simplification in paragraph three. From now on, I will consider a language to be a set (finite or infinite) of sentences, each finite in length….The fundamental aim in the linguistic analysis of a language L is to separate the grammatical sequences which are the sentences of L from the ungrammatical sequences which are not sentences of L and to study the structure of the grammatical sequences. The grammar of L will thus be a device that generates all of the grammatic sequences of L and none of the ungrammatical ones (p. 13, his italics, my bold). But it helps to distinguish grammars from the aims of linguistic analysis, and to recall the first paragraph. Syntax is the study of the principles and processes by which sentences are constructed in particular languages. Syntactic investigation of a given language has as its goal the construction of a grammar that can be viewed as a device of some sort for producing the sentences of the language under analysis. More generally, linguists must be concerned with the problem of determining the fundamental underlying properties of successful grammars. The ultimate outcome of these investigations should be a theory of linguistic structure in which the descriptive devices utilized in particular grammars are presented and studied abstractly, with no specific reference to particular languages. One function of this theory is to provide a general method for selecting a grammar for each language, given a corpus of sentences of this language (Chomsky 1957, p. 11). The talk of languages as sets, which Chomsky elsewhere eschews, was presumably to ease MIT students into the subject via his subsequent discussion of kinds of recursion.
generates usable \(<\mu, \pi>\) pairs is hard enough; and a natural implementation of such a procedure may connect “slightly foreign” pronunciations with meanings, in ways that are felt to be degraded, without representing the degree to which such connections deviate from some ideal.\(^{17}\)

2.2 An Extensional Alternative

At this point, I would like to think that when I talk about human linguistic meanings, it is clear what I am talking about. But some specialists will want to know why I did not begin with Lewis’ (1975) conception of meanings and human languages; and it is instructive to contrast his starting point with mine. Readers who want to move on can safely skim the rest of this section.

Lewis begins with an extensional conception of languages.

What is a language? Something which assigns meanings to certain strings of types of sounds or marks. It could therefore be a function, a set of ordered pairs of strings and meanings (p.3).”

His opening question, fit for a Platonic dialogue, seems to call for a general conception of what a language is; and the sentence fragment that follows sounds suitably generic. Socrates might have interjected to slow things down. (Does a language assign strings to meanings? If a coin depicts a face on one side and a building on the other, does the coin assign the face to the building?) But niceties aside, Lewis’ initial characterization of languages is fairly innocuous. His next sentence, however, begins with ‘It could’. The apparently generic has become rather specific, as in (35).

(35) Something which is bad for bees got into the hive. It could be a fungus.

Presumably, Lewis was not saying that even given a generous conception of languages, it could be that all of them are sets. If we can invent I-languages that assign meanings to strings, then at least some “meaning assigners” are procedures rather than functions in extension. And as the rest of Lewis’ essay makes clear, he wanted to provide a description of human languages that coheres with his answer to the mass-noun version of his initial question.

What is language? A social phenomenon which is part of the natural history of human beings; a sphere of human action wherein people respond by thought or action to the sounds or marks which they observe to have been so produced (p.3).

My preference would be to replace Lewis’ talk of language with talk of socially regulated uses of I-languages. But in any case, his proposal was that human languages are sets that are related to social phenomena in ways that he describes by appealing to conventions of “truthfulness and trust.” I’ll return to his emphasis on the conventionality of language (use). Though with regard to first quoted passage above, one wonders why a human language “could therefore be” a function-in-extension just because it assigns meanings to certain strings.\(^{18}\)

---

17 The deviance might be registered by other systems. But specifying the alleged set of ordered triples in a way that accommodates (32) requires specification of a procedure that connects the meaning of (32a) with both the pronunciation of (32) and a number that reflects the relevant degree of deviance. In terms of Marr’s (1982) levels of analysis, one might describe an acquired I-language as an implemented procedure that maps certain meanings—generable from a stock of lexical meanings via some procedure—onto pronunciations in a roughly “reversibility” way that makes it possible to compute the one or more meanings of each pronunciation. But one doesn’t need to talk about extensions to distinguish the question of what a system computes from the question of how it computes outputs given inputs. Though it may be relevant that while lexical homophony is common, lexical synonymy is not. There may be no “intra-language” lexical synonymy apart from a few special cases—e.g., ‘lawyer’ and ‘attorney’ in American dialect—as children assume that distinct lexical pronunciations signal distinct meanings.

18 As we’ll see, Lewis’ claim was not merely that there are possible worlds in which certain sets are the counterparts of human languages. His ‘could’ is not that weak. Though it may be useful to imagine a paper that begins as follows: What is a language? Something that connects meanings with pronunciations. So it might be an acquirable procedure that can be used in communication. What might a sentence meaning be? Something that often interfaces with cognitive systems to a yield thought. It could therefore be an instruction to build a thought.
Lewis goes on to say “if σ is the domain of a language £, let us call σ a sentence of £.” But he was not introducing ‘sentence’ as a technical term that applies to any string to which a language assigns a meaning. Lewis was using a more ordinary notion of sentence and initially restricting his notion of a meaningful expression to sentences. He later introduces the idea of a grammar characterized in terms of finitely many atomic strings (lexical items) and combinatorial operations. But at the outset, he offers a third question-answer pair.

What could a meaning of a sentence be? Something which, when combined with factual information about the world—or factual information about any possible world—yields a truth value. It could therefore be a function from worlds to truth-values (p. 3).

Lewis assumes that we know what he is talking about when he talks about sentence meanings. His question invites a proposal; and the ensuing sentence fragment indicates the kind of answer he has mind. The suggestion that meanings are functions is even more specific and controversial.

On first reading, it can seem that Lewis started with platitudes. But in my view, he mixed stipulations and substantive assumptions. His claim that human languages are sets is part of a package that includes tendentious claims about sentence meanings and how they are related to conventions. Put another way, it turns out that his conception of human languages is motivated largely by his account of sentence meanings and how sentences come to have such meanings. Not that there’s anything wrong with that. To use a contemporary idiom, hypotheses concerning the nature of human languages should be evaluated in light of our best accounts of semantics and metasemantics. But our best conception of human languages should also inform our accounts of meaning. And I see no good independent reasons for thinking that these languages are sets.

In this context, it is important to be clear that various I-languages can be described as English dialects—or as dialects “of English”—without English being any one of the acquirable human languages. Prima facie, there are many ways to be a speaker of English: American, Australian, British, and Canadian ways; young child ways, professional scientist ways; etc. Being a speaker of an English dialect seems to be a multiply realizable property whose instances are similar in ways that matter for certain practical purposes. (Likewise, there are many ways to be a speaker of a Germanic language.) So we can use ‘English’ to talk about a complex respect in which certain individuals are similar, even if there is no English language that each of these individuals have imperfectly acquired; ep. Dummett (1986). In ordinary contexts, we speak of English and French as languages. But when theorizing, different standards may apply. One can insist that if Brit and Candice speak English, then there is something that Brit and Candice speak. But then the “thing spoken,” in the insisted sense, may be a vague class of I-languages such that Brit and Candice each acquired at least one of them.¹⁵

Mere talk of English does not ensure that any I-language, much less any set of sentences, is common to all speakers of English. We can focus on a particular dialect, INGLISH. And for these purposes, let’s waive concerns about examples like (32) and (34).

(32) *the child seems sleeping
(34) *the lawn needs mowed

To focus on a set, I think we need to assume that the relevant speakers (e.g., BBC news readers) have a procedure, i–INGLISH, that determines an extension. One can hypothesize that some set

¹⁵ English dialects may be characterized in terms of “central cases” and a notion of mutual intelligibility that need not be transitive; think Brooklyn and Glasgow. One can stipulate that people “share a language” if and only if they can communicate linguistically. But then positing shared languages does not explain successful communication.
is the set of **INGLISH** sentences; call this alleged set **e-** **INGLISH**. But unless it is characterized as the extension of a certain procedure, I don’t know which set **e-** **INGLISH** is supposed to be, much less how “it” might be related to meanings or episodes of acquiring **INGLISH**.

I can consider many sentences that are allegedly elements of **e-** **INGLISH**. But there are boundlessly many sentences of **INGLISH** that I have never considered. So I need some non-enumerative way of thinking about them, in order to assess Lewis’ proposal. One might describe **e-** **INGLISH** as the set of strings that a monolingual BBC news reader *could* understand as sentences, given a certain dictionary and a suitably idealized sense of ‘could’. But this seems to presuppose that a competent speaker of **INGLISH** has an expression-generating *procedure* whose lexicon can be expanded. So again, it seems that any characterization of the alleged set relies a prior notion of the language as intensional in the Church/Chomsky sense.

Moreover, if languages are sets, then adding a lexical item makes for a distinct language in the same sense as adding a composition principle. From a set-theoretic perspective, the effect is the same—viz., a “larger” language with more expressions. I readily grant that adding lexical items changes a language. But one can think of a human language as a generative procedure that is partly characterized by a certain list of atomic expressions, and partly characterized by more general procedures for creating and extending such lists. By contrast, individuating human languages as sets of expressions seems like individuating animals as sets of molecules and insisting that growth be described as replacing one animal with another.

Correlatively, a set-theoretic perspective trivializes the differences between languages governed by different compositional principles. One can follow Lewis in saying that any such difference is conventional in his technical sense. But this dilutes the sense in which familiar lexical differences, like American vs. English ‘biscuit’, *are* conventional. Worse, even if it is a matter of convention that children acquire **e-** **INGLISH** as opposed to other sets, that leaves us the question of why children acquire **e-** **INGLISH** as opposed to “larger” variants that allow for more compositional homophony of the sort described in section one; cp. Chomsky (@RefOnLg).20

One can say that this psychological question has no bearing on what human languages are. But this hypothesis requires defense. Alternatively, one can stipulate that one is talking about certain sets, leaving it open how they are related to human language acquisition; but then one has to say which sets one is talking about. In any case, theorists cannot stipulate that human languages are not languages of a certain *kind* whose instances are child-acquirable *procedures* that connect meanings with pronunciations in certain ways. On the contrary, this conception of human languages seems to be independently plausible.

This is not to deny the relevance of communication in language acquisition. A child may initially use an I-language in a way that leads to correction or failure of communication; and this may lead the child to adopt a slightly different I-language. But children regularly go through

---

20 See also Baker’s (2001) discussion of Navajo and Japanese. Lewis does not consider constraints on ambiguity. But he replies to the objection that “Language is not conventional. We have found that human capacities for language acquisition are highly specific and dictate the form of any language that humans can learn and use (p.24).” It may be that there is less conventionality than we used to think: fewer features of language which depend on convention, more which are determined by our innate capacities and therefore common to all languages which are genuine alternatives to our actual language. But there are still conventions of language...That is established by the diversity of actual languages. There are conventions of language so long as the regularity of truthfulness in a given language has even a single alternative (24).

Perhaps parameters of grammatical variation can be described in terms of Lewisian conventions regarding how the parameter values are set. But a “regularity of truthfulness” may not play any explanatory role here.
stages of employing generative procedures that differ (extensionally and intensionally) from those of local adults.\textsuperscript{21} Cases of creolization and language change are dramatic instances in which a group of children stabilize on a language other than the one(s) used by local adults. Though even bracketing such cases, it is hard to see how any group of humans could determine an unbounded set of sentences unless individuals acquire procedures that generate the sentences.

Lewis grants this last point when he introduces a technical notion of grammar to describe the meanings of non-sentential expressions. He says, citing Chomsky (1965) and others, that a set of sentences can be described as the product of a finite lexicon and a generative procedure. This is how Lewis tries to capture the idea that sentence meanings are determined by lexical meanings and grammatical structure: a grammar describes the sentences of a language as having certain constituency structures; it assigns meanings of some kind—e.g., functions from possible worlds to sets—to the atomic constituents; and it specifies the relevant sentence meanings compositionally, along the lines suggested in Lewis (1972). But instead of concluding that human languages are grammars/procedures, Lewis says that “a grammar, like a language, is a set-theoretic entity which can be described in complete abstraction from human affairs (p. 19).”

He defines grammars so that each “uniquely determines the language it generates,” but a language “does not uniquely determine the grammar that generates it (p.18).” His grammars are more abstracted from language than languages are. But even if there is a Lewisian grammar for \textit{e-INGLISH}, this does not explain how the relevant sentence meanings are related to lexical meanings, much less how these sentence meanings are not pronounced. At best, such a grammar indicates how a certain kind of mind might abstract lexical meanings from sentence meanings, given hypotheses about the relevant constituency structures and composition principles; see chapter two. So why define grammars this way instead of saying that human languages are internalized grammars/procedures in Chomsky’s sense?

\subsection*{2.3 Grammars, Meanings and Sentences}

Lewis’ response is indirect but again instructive when thinking about potential starting points for discussions of linguistic meaning. I think many debates in philosophy of language are rooted in disagreements or unclarity about whether human languages are expression-generating procedures as opposed to sets of expressions. Lewis was clear, making it easy to isolate the disagreement. He also considered one way of individuating languages more finely than he recommended.

Different grammar, different language—at least if we ignore superficial differences between grammars. Verbal disagreement aside, the place I gave to my so-called languages ought to have been given to my so-called grammars. Why not begin by saying what it is for a grammar \( \Gamma \) to be used by a population \( \mathcal{P} \)?

To this objection, Lewis offered a reply.

Unfortunately, I know of no promising way to make objective sense of the assertion that a grammar \( \Gamma \) is used by a population \( \mathcal{P} \), whereas another grammar \( \Gamma' \), which generates the same language as \( \Gamma \), is not. I have tried to say how there are facts about \( \mathcal{P} \) which objectively select the languages used by \( \mathcal{P} \). I am not sure there are facts about \( \mathcal{P} \) which objectively select privileged grammars for those languages...a convention of truthfulness and trust in \( \Gamma \) will also be a convention of truthfulness and trust in \( \Gamma' \) whenever \( \Gamma \) and \( \Gamma' \) generate the same language (p. 20).

\textsuperscript{21} Thornton (1990, 1996) offers striking examples of children using question forms that are not attested in the “local” language, but still subject to constraints of universal grammar. She thereby provides interesting support for the “continuity hypothesis,” according to which language acquisition is effectively a process of “trying out” a series of I-languages; see Pinker (1994), Crain (1991), Crain and Pietroski (2001, 2011).
Given how Lewis defined grammars, extensionally equivalent grammars may be equally good (and so equally bad), equally selected by populations, etc. But this may be a symptom of unfruitful definition. One wants to know what gets explained by appealing to Lewisian grammars of E-languages and identifying human languages with the latter. Individuals can implement a certain procedure without implementing extensionally equivalent procedures. The I-languages used by a population are those used by the relevant individuals. And even if populations select Lewisian languages, this may require children who acquire I-languages.

According to Lewis, £ is the language used in a population P “by virtue of the conventions of language prevailing in P (p. 7);” where conventions, in his technical sense, are special cases of mutually recognized regularities of action/belief in a population of individuals who can act rationally. More specifically, he says that a convention of “truth and trustfulness,” sustained by “an interest in communication,” is what makes a particular set of sentences the language of a given population (p. 10). Members of P conform to such a convention by trying “never to utter any sentences of L that are not true in £” and trying “to impute truthfulness in £ to others,” thereby tending to “respond to another’s utterance of any sentence of £ by coming to believe that the uttered sentence is true in L’” (p. 7). But even if human beings approximate this ideal, I don’t see how an actual person could use or hear (1) as a true sentence—

(1) the guest is easy to please
at least not in a systematic way that extends to (2) and other examples—

(2) the guest is eager to please
without implementing a suitably constrained generative procedure. Nor do I see how there could come to be a convention of connecting certain pronunciations with their alleged truth conditions unless the relevant individuals already understood the pronunciation in a common way.

I grant that among speakers of English dialects, the pronunciation kao is conventionally associated with cows, as opposed to horses or things that are easy to do; see §1.2 above. Though like Davidson (1986), I suspect that speakers can connect meanings with pronunciations in unconventional ways and still be understood. Moreover, we hear (36) as a sentence,

(36) all mimsy are the borogoves
rather than a mere string of three words and two non-words. Competent speakers of English know that ‘mimsy’ is an adjective, and that ‘borogoves’ is a plural (and hence count) noun; see, e.g., Higginbotham (198x). In this sense, we understand (36) as we understand (37),

(37) all tangled are the cypresses
except for not knowing which adjectival or count-noun meanings ‘mimsy’ and ‘borogove’ have.

It is conventional/arbryitary that the pronunciations of ‘all’, ‘the’ and the English plural marker are connected to the meanings of certain closed class (functional) expressions, which provide useful markers for open class expressions. But to learn the conventional aspects of what ‘tangle’ and ‘cypress’ means, or even formulate a hypothesis about these aspects of meaning, one may need an I-language that imposes constraints on the kind of meaning that lexical items of a given type can have. I will argue for a view according to which these constraints are severe. If the conventionality of meaning is thus limited to certain aspects of lexical meaning, and perhaps a few dimensions of variation regarding how structured expressions are pronounced (e.g., whether heads begin or end their phrases), it can be accommodated in various ways.

Suppose that the meaning of ‘cypress’ is an instruction for how to access a concept of a

---

22 If a grammar Γ for a human language £ connects sentences of £ with (specifications of) truth conditions that are established by virtue of how sentences of £ are used in communication, then any notion of sharing Γ may well depend on a more notion of a community using the relevant sentences in ways that establishes their truth conditions.
certain sort from a lexical address that was created and linked to the pronunciation *saipras* when the lexical item was acquired. A child may create such an address without having any concept that is specifically about cypresses (as opposed to beeches, elms, yaks, or zithers). So the meaning of ‘cypress’ can differ from the content of any concept, even if there is a convention of using *saipras* to refer to cypresses, perhaps allowing for “deferential” concepts of whatever certain experts think about when using a concept they have linked to *saipras*; cp. Putnam (1975). To assume otherwise is to start with a substantive proposal about what meanings are.

Lewis also assumes that sentences are privileged expressions that exhibit a special kind of meaning, from which lexical meanings can somehow be abstracted. But human languages may not even generate sentences as such. As discussed in chapters one and six, the expressions we call sentences may be grammatically varied instructions for how to build mental predicates that are easily converted into *thoughts* that exhibit the kind of subject-predicate structure discussed by classical logicians and grammarians. It may be that certain *mental sentences* can be described with “rewrite rules” that include those in (38);

(38) \[ \text{Sentence } \Rightarrow \text{Subject copula Predicate} \]
\[ \quad \text{Subject } \Rightarrow \text{Quantifier Predicate} \]
\[ \quad \text{Subject } \Rightarrow \text{Name} \]
\[ \quad \text{Predicate } \Rightarrow \text{Predicate Predicate} \]

where copulas, quantifiers, names, and predicates are mental symbols of certain kinds, and predicates can be complex symbols—mental analogs of ‘brown cow’—that have predicates as constituents. Chomsky (1957, 1965) noted that given a grammar for English that employs rewrite rules like “S \( \Rightarrow \) NP (aux) VP,” one might say that an “NP of S” is a subject of a sentence. But he went on to argue that ‘subject’ does belong in the vocabulary of a good grammatical theory; human languages do not generate subjects as such.

Linguists have since replaced ‘S’ with a bundle of phrase-like projections of functional items that include tense and agreement morphemes along with a variety of complementizers. This raises questions about what sentences are, and whether it reasonable to assume that there is a grammatical notion of sentence corresponding to the notion of a truth-evaluable thought. But theories of grammatical structure—and to that extent, theories of the expressions that human languages generate—have been improved by *not* positing a special category of sentence. So while such a category often plays a special role in the stipulations regarding invented formal languages, grammatical structure may be independent of any notion of sentence.

By contrast, for Lewis, sentences are prior to grammars. While granting that we should not “discard” notions like phrasal meaning (relative to a population P), or the “fine structure of meaning in P of a sentence,” Lewis says that these notions “depend on our methods of evaluating grammars.” In his view, a grammar \( \Gamma \) is used by P if and only if \( \Gamma \) is a best grammar for a language \( \mathcal{L} \) that is used by P in virtue of a convention in P of “truth and trustfulness in \( \mathcal{L} \).” One might have thought that a “best” grammar for the alleged set \( \mathcal{L} \) would be one that best depicts the procedures acquired by members of P. But according to Lewis, “it makes sense to say that languages might be used by populations even if there were no internally represented grammars.”

---

He adds that “the point is not to refrain from ever saying anything that depends on the evaluation of grammars. The point is to do so only when we must, and that is why I have concentrated on languages rather than grammars” (pp. 20-21). But in reply to a worry that he is needlessly hyopstatizing meanings, he says “There is no point in being a part-time nominalist. I am persuaded on independent grounds that I ought to believe in possible worlds and possible beings therein, and that I ought to believe in sets of things I believe in (p. 17).” So why be a part-time grammarist, given Chomsky’s (1957, 1964, 1965) reasons for thinking that children acquire generative procedures?
I can tentatively agree that £ is used by P if and only if everyone in P possesses an internal representation of a grammar for £, if that is offered as a scientific hypothesis. But I cannot accept it as any sort of analysis of “£ is used by P”, since the analysandum clearly could be true although the analysans was false (p. 22).

Note the shift from his opening question—what is a language?—to a search for a certain kind of analysis of what it is for a language to be used by a population. This shift is prompted by the stipulations that languages are set of sentences, and that grammars connect these sentences conventionally determined contents, along with a willingness to accept the consequences for how grammars are related to non-sentential expressions. I think this ordering of priorities, often presupposed, should be reversed. In my view, grammars/languages are fundamentally concerned with lexical items and how they can be combined to form phrases, which can used to construct mental predicates that can figure in sentences.

Since a Lewisian grammar is even more abstracted from language use than sets of sentences are, while meanings are still being characterized in terms of truth, Quine’s (1960) worries about the “indeterminacy” of meaning are not far away. Lewis’ talk of evaluating grammars reflects this. But he does not engage with Chomsky’s (1965) notion of an evaluation metric or the correlative notion of “explanatory adequacy,” which is rooted in the idea that children somehow use experience to select particular human languages from the space of possible human languages. Chomsky stressed the importance of the theoretical vocabulary that linguists use to specify expressions and their meanings. He urged the project of formulating a grammar for each human language in a vocabulary that makes it possible to (i) formulate a grammar for all and only the expression-generating procedures that children can naturally acquire, and (ii) show how evidence available to a child who acquires one of these procedures could allow for selection of that procedure; see Pietroski (2015) for discussion.

By contrast, Lewis speaks of populations selecting languages by virtue of using sentences in rational ways. On his view, if language use is not a “rational activity” for young children, they are not “party to conventions of language” or “normal members of a language-using population.” Perhaps language is first acquired and afterward becomes conventional...I am not concerned with the way in which language is acquired, only with the condition of a normal member of a language-using population when he is done acquiring language (p.26).

I’m not sure what it is to acquire language in Lewis’s sense. But even if one wants to focus on the use of human languages, being unconcerned with their acquisition may not be a viable option. The languages we naturally use are the ones that children naturally acquire; and these languages seem to be procedures that connect meanings with pronunciations in certain ways.

Indeed, children can acquire a procedure that connects several meanings with the pronunciation of (39) but only one meaning with the pronunciation of (40); cp. (14-15).

(39) we saw the jabberwock whiffling towards the borogove
(40) this is the borogove that we saw the jabberwock whiffling towards

So perhaps we should take this as indication of what meanings are, instead of disconnecting meaning from language acquisition and insisting that meanings are Lewisian contents.24

---

24 If sentences are pronounceable instructions for to how build thoughts, there is no guarantee that an instruction like (45) can be executed, much less that executing the instruction yields a thought with a Lewisian content. We may be able to create a lexical address for ‘whiffle’, even if there is no concept of whiffling to store at that address; and an instruction may call, in vain, for a concept from that address.
2.4 Back to Procedures

Lewis started with the idea that sentence meanings can be characterized in terms of truth, and that the source of meaning is convention. This led him to adopt an extensional conception of human languages. I think we should start with the idea that human languages are procedures that connect meanings with pronunciations in certain ways, even if this leads us to reject Lewis’ conception of meaning. In my view, it is no accident that early advocates of Truth Conditional Semantics did not take themselves to be offering theories of biologically implemented expression-generating procedures. I focused on Lewis because he was unusually explicit about what he took human languages to be. But his audience was prepared for the suggestion that a sentence meaning is “Something which, when combined with factual information about the world—or factual information about any possible world—yields a truth value (p.3).”

Frege, Russell, and Wittgenstein had suggested various ways in which sentences of an ideal language could have truth conditions. As discussed in chapter two, Tarski and Church showed how to encode the suggestions as explicit algorithms that connected unboundedly many well-formed formulae of an invented language with (conditional specifications of) truth values—or with propositions that somehow determine truth values relative to possible worlds. Amidst other excitement in the late 1960s, Davidson then suggested that human language sentences also have truth conditions, albeit in context-sensitive ways that are far from obvious. Davidson was not explicit about what he took human languages to be; and as discussed below, he tended to ascribe truth to utterances of sentences, as opposed to linguistic expressions themselves. But he and Lewis seemed to a share (perhaps via Quine and Wittgenstein) the idea that languages are abstractions, and that linguistic usage is somehow more fundamental: theorists can and perhaps should appeal to sets of expressions in order to capture regularities in usage; but positing internalized languages/grammars, and saying that usage is a product of these and other mental entities, is a rash hypothesis whose relation to meaning is unclear.

From this perspective, characterizing meaning in terms of truth is a way of suggesting that any semantic regularities in usage can be captured in truth-theoretic terms, without positing independent meanings that allegedly determine the truth conditions of sentences. This suggestion may even be correct, given a suitably construal of “semantic regularities in usage.” But in my view, the phenomenon of constrained homophony makes it clear that particular uses of human languages are just that: uses of languages that are independent of (and prior to) any particular behavior. And if a conception of meaning is plausible only on the assumption that human languages are sets of expressions, as opposed to expression-generating procedures, then so much the worse for that conception. If the task is to say what human linguistic meanings are—as opposed to saying what the analogs of meanings would be if the only languages were those that conformed to certain idealized conceptions of communication—then we can’t ignore the reasons for taking human languages to be expression-generating procedures that connect meanings with pronunciations in certain ways but not others. Whatever human linguistic meanings are, human I-languages connect them with pronunciations in accord with certain constraints.

Following Higginbotham (1985) and others, one can hypothesize that human I-languages connect the pronunciations of declarative sentences with mappings from contexts to truth conditions, thereby (and more generally) connecting reference/satisfaction conditions with pronunciations. I take this proposal seriously, and will offer reasons for rejecting it in favor of the idea that human I-languages generate pronounceable instructions for how to build concepts of a special sort. But I see no point in arguing about the alleged meanings of alleged E-language analogs of the languages that human children actually acquire.
3. What’s Truth Got to Do with It?

Davidson (1967b) conjectured that for each human language H, the sentences of H have truth conditions that can be described (a la Tarski) with a finitely statable theory of truth that can serve as the core of a good theory of meaning for H. Given how influential this proposal has become, one can forget how initially implausible it is, and how tendentious it is to assume that ordinary sentences have truth conditions. So let me offer some reminders, while also making it clear that my initial conception of meaning leaves room for both accepting and rejecting “truth conditional semantics.” Chapters three and four will develop the point that Davidson’s conjecture should be assessed in light of the more plausible—and historically prior—claim that human languages are naturally acquired expression-generating procedures.²⁵

3.1 Davidson’s Bold Conjecture

Tradition held that pronounceable sentences have meanings that are somehow akin to thoughts, and that truth is property of certain thoughts, not pronounceable sentences. If only because humans can think about numbers and geometric figures, it seems obvious that whatever thoughts are, they are individuated more finely than sets of potential situations (or worlds, or contexts, or mappings from contexts to worlds). So if a linguistic expression Σ determines some such set, one might think this is because μ(Σ)—the meaning of Σ—determines a thought that determines the set. It is even more obvious that in any context, many thoughts have the same truth value. So if μ(Σ) “yields a truth value” when “combined with factual information,” one might think this is because μ(Σ) somehow determines a thought that determines the truth value given the facts.

Prima facie, trying to characterize meanings in terms of truth is like trying to characterize genomes in terms of mappings from environments to organisms. Cutting out middlemen is not always a good idea. In my view, meanings can be used to assemble concepts, which can be used to make judgments that are true; but meaning is not connected to truth in any direct way that makes it possible to characterize meaning in terms of truth. Davidson (1967a, 1967b, 1968) and others, however, have advocated versions of thesis (D).

²⁵ In terms of the history, Chomsky (1957, 1964, 1965) did not use the term ‘I-language’. But he clearly held that human languages are generative procedures. Lewis (1975) explained why he nonetheless took languages to be sets; see section two. Davidson (1965, 1967a, 1967b) was less explicit about what he took languages to be, and how he thought truth theories were related to the competence that Chomsky was trying to describe; though see Davidson (1967b, note 14) and Davidson (1986). It is also worth noting that an earlier version of Lewis (1975) was prepared in 1968 and published in Italian as Lewis (1973). Lewis’ conception of grammar may also have been influenced by Montague, who regarded “the construction of a theory of truth—or rather, of the more general notion of truth under an arbitrary interpretation—as the basic goal of serious syntax and semantics” (Montague 1970, p.xx, my italics). Montague held that “developments emanating from the Massachusetts Institute of Technology offer little promise towards that end;” though he did not discuss the facts that motivated Chomsky’s conception of syntax. And while Montague claimed to be like Davidson in this respect, Davidson hoped for a “rapprochement” between his project and Chomsky’s; cp. Partee’s (1975, 197x) attempts to reconcile Montague’s conception of semantics with a more plausible conception of human language syntax. Harman (1974) offered apt cautionary notes and a forceful critique—in my view, still unrebutted—of the idea that truth theories can serve as theories of meaning for human languages; see also Chomsky (1977). Lewis (1972) had said, “Semantics with no truth conditions is no semantics;” but as Harman noted, this was effectively a stipulation about how to use ‘semantics’, not a motivated criticism of Katz and Fodor’s (1963) psychologistic conception of meaning. Evans (1982), Higginbotham (1985), and others tried to explicitly connect Davidson’s program with Chomsky’s. See Larson and Segal (1995) for a systematic presentation. Heim and Kratzer (1998) likewise connect Montague’s program with Chomsky’s. But there is a counterfactual “history of semantics” in which central insights from the 1960s and 1970s were developed by focusing on questions concerning how generable linguistic expressions interface with human concepts, as opposed to questions about how such expressions/concepts are related to truth; cp. Bach (197x).
(D) for each human language $H$, there is a theory of truth that is the core of a correct theory of meaning for $H$.

Given how mainstream this idea has become, I don’t want to start with a notion of meaning that precludes (D). But as we’ll see, advocates of (D) can and should grant that human I-languages connect meanings with pronunciations. The thesis itself, however, calls for explanation.

For these purposes, a theory is a finite list of axioms and rules for deriving theorems. A theory of truth for a language is a theory such that for each sentence $S$ of the language, there is a theorem of the form ‘True($S$, $c$) $\equiv$ $\Psi(c)$’; where ‘$c$’ ranges over contexts of some kind, and ‘$\equiv$’ is the material biconditional. Tarski (1933, 1944) showed how to provide truth theories for certain invented languages that can be viewed as special cases, with vacuous relativization to contexts. As reviewed in chapter two, each sentence of a Tarskian language corresponds to a theorem of the form ‘True($S$, $c$) $\equiv$ $p$’; where ‘$p$’ can be the form ‘for each sequence $\sigma$ of potential values for any variables, $\Psi(\sigma)$’. As discussed in chapter four, there can also be truth theories for languages that have indexical and deictic expressions, like ‘I’ and ‘this’, along with grammatical moods (e.g., declarative, interrogative, and imperative) that are related to kinds of speech acts (e.g., assertions, queries, and requests) in meaningful ways that can be captured by supplementing a truth theory with claims about how expressions are naturally used.

A truth theory can thus be the core of a proposal regarding both the context-invariant “semantic character” that expressions of a given language exhibit, and at least some of the ways in which this character is related to the natural use of these expressions in particular contexts. Though not every truth theory is suited to being the core of a meaning theory. For example, a truth theory for $H$ might have theorems of the form ‘True($S$, $c$) $\equiv$ Expresses($S$, $T$) $\&$ True($T$, $c$)’; where ‘$T$’ ranges over representations of some other (perhaps mental) language $L$, and ‘Expresses($S$, $T$)’ is a variant of ‘$\mu(S) = T$’. Such a theory doesn’t explain meaning in terms of truth, in any interesting sense, even if it does explain what it is for a sentence of $H$ to be true.

One can, however, invent languages whose signals are connected with—and only with—the interpretations that are directly specified by truth theories according to which sentences have context-sensitive truth conditions. Such languages can provide useful models of human languages. Davidson speculated that children acquire languages of this sort. From a traditional perspective, this might seem like a rejection of the idea that linguistic expressions have meanings. But from the perspective urged in section one, Davidson offered a specific proposal about the meanings that human languages connect with pronunciations: the meaning of a sentence, in effect, the set of contexts that makes it true; where contexts may turn out be nothing more than assignments of values to context-sensitive aspects of sentences.

Lewis offered a slightly different version of (D) according to which meanings are conventionally determined sets of logically possible worlds. Indeed, (D) is abstract enough to permit many variants. But in thinking about truth-conditional semantics, as a program that can be developed in many ways, it can be helpful to focus on two subsidiary theses.

(D1) for each human language $H$, there is a correct theory of truth for $H$.

(D2) a suitably formulated theory of truth for a human language $H$ can be the core of a correct theory of meaning for $H$.

Each of these claims is hard to defend; and their conjunction is still weaker than (D). In §3.2, I focus on some difficulties for (D1). Then I’ll turn to (D2), some motivations for (D) despite the

---

26 Kaplan (1978b, 1989) stressed the importance of evaluating a sentence relative to a context, and asking at which worlds the evaluated sentence is true; see Stanley (200x) for excellent discussion in a broader historical context.
obvious objections, and some advertisements of why I think these motivations are insufficient given the objections. But my point is emphatically not that thesis (D) has been unhelpful.

On the contrary, I take it as obvious that a great deal of insightful work has been motivated by this framework assumption and presented in terms of it. Nonetheless, I think that (D) is like most good proposals: helpful but false, illuminating and yet inadequate. I would be delighted if it turns out that this book includes any such proposal. Thesis (D) reflects one of the first serious attempts to describe human linguistic meaning as *compositional*, and perhaps the first that did simply blame the compositionality of meaning on something else—e.g., the compositionality of thought. Those of us who get think about these topics after Davidson have a great advantage those who came before. But first attempts are never right. The job is to do better.

### 3.2 Theories of Truth for Human Languages?

Defending (D1) is an enormous task given the many human language constructions, only some of which have analogs in invented languages. Moreover, the inventions were designed to idealize away from certain features of the expressions that children naturally acquire.

For example, Tarski himself thought that examples like (41)

> (41) The forty-first numbered example in this chapter is not true.

showed that (D1) is hopeless unless restricted to certain *fragments* of human languages that do not include semantic vocabulary and the resources needed to designate sentences of such languages. With some exceptions noted in chapter four, advocates of truth conditional semantics have often set “Liar Sentences” aside as special case, as if examples like (41) present general puzzles about truth—which logicians should deal with—but not direct challenges to the idea sentences of a human language are among the things that can be true or false, at least in suitable contexts. Often, the assumption seems to be that some correct response to the paradoxes can be accommodated with at most minor tweaks to extant versions of truth conditional semantics. But as we’ll see in chapter four, Liar Sentences suggest simple (and in my view, strong) arguments that no expressions generated by a human I-language are among the truth-gradable things. Such arguments are surely relevant to evaluating (D1), even if there are potential replies. Moreover, the currently available replies involve stipulations regarding truth that seem to be at odds with (D2). In short, Liar Sentences present a serious challenge to (D), not a technical difficulty that can be bracketed as a “philosophical” problem.

One cannot complain that examples like (42) have been ignored.

> (42) Some ancient astronomer wondered whether Hesperus is Phosphorus

But much of the attention has focused on the phenomenon of “referential opacity;” replacing one name for Venus with another does not always preserve truth. This does present a technical challenge for advocates of (D1). However, I think “attitude reports” illustrate a deeper reason for thinking that (D) is false, especially given examples like (41). And it is worth spelling this out.

The technical challenge is to specify suitable “semantic values” for verbs like ‘wondered’ and complementizer phrases like ‘whether Hesperus is Phosphorus’, while also providing a suitable composition rule that determines plausible semantic values for the corresponding verb phrases. Ignoring the interrogative aspect of ‘wondered’, one might focus on (43-46).

> (43) Sam denied that Hesperus is Phosphorus.
> (44) Sam denied that Hesperus is Hesperus.
> (45) Sam denied that Hesperus is often visible shortly before sunrise.
> (46) Sam denied that Hesperus is often visible shortly after sunset.

Suppose that Sam was a reasonable ancient astronomer who had often seen Hesperus shortly after sunset, was certain that Hesperus is self-identical, but rejected the new-fangled
theory that the evening star is also be the morning star. Then there could be a scenario in which (43) and (45) are used to correctly report some of what Sam did, in the course of a certain conversation, while (44) and (46) would not be so used in that scenario. And if Sam’s student claimed that Hesperus would be visible the next morning, while Sam’s teacher claimed that Hesperus would be visible on a future evening, then a reporter could use (47) but not (48)

(47) Sam rejected his student’s claim.
(48) Sam rejected his teacher’s claim.

to correctly report something Sam did. So at least prima facie, (D1) implies that a correct theory of truth for English will connect the direct objects of (47) and (48) with different claims, at least relative to the context in question. If ‘claim’ is not true of both a claim by the student and a distinct claim by the teacher, it’s hard to see how a theory could plausibly assign truth conditions to (47) and (48). Likewise, it seems that a correct theory of truth for English will have to connect the ‘that’-clauses in (43-46) with distinct entities of some kind.\(^{27}\)

There have been many proposals regarding the entities that ‘that’-clauses and words like ‘claim’ (‘belief’, ‘report’, etc.) are allegedly true of. Theorists have invoked Fregian senses, sets of logically possible worlds, linguistic expressions, hybrids of these, and other still more exotic abstracta. To take the most simple-minded idea, one might say that ‘that Hesperus is Phosphorus’ is true of the sentence ‘Hesperus is Phosphorus’, which differs from ‘Hesperus is Hesperus’; cp. Carnap (1955). If we take sentences to be meaning-pronunciation pairs, then by hypothesis, the semantic values of ‘that’-clauses are individuated as finely as such pairs. But that’s still not finely enough, given examples like (49) and (50); cp. Kripke (1979a).

(49) Sam said that I am an idiot.
(50) Pierre said that Paderewski is a great musician.

It matters whether (49) is uttered by Sam’s student or his teacher, and whether (50) is uttered by someone who knows that Pierre mistakenly thinks there are two Paderewskis. But one might think that such examples can be accommodated by taking ‘that’-clauses to be true of contextually evaluated sentences, at least if proper names are complex expressions that include a demonstrative index; see, e.g., Larson and Ludlow (1993) and references there. Though even if the technicalities can be worked out, this still wouldn’t show that any such account is plausible. On the contrary, any such theory leaves us with the question of whether its theorems are true.

Consider the “hypothesis” that (43) is true, relative to context c,
(43) Sam denied that Hesperus is Phosphorus.

if and only if Sam denied ‘Hesperus is Phosphorus’ as relativized to c. To make a judgment about whether this biconditional is plausible, we need to know what it is for someone to deny a contextually relativized sentence of English; and as Kripke showed, the first (quasi-behavioristic) suggestions that come to mind lead to contradiction given plausible assumptions. Compare the

\(^{27}\) Or employ a composition rule (for combining a verb like ‘denied’ with a ‘that’-clause) that quantifies over such entities; cp. Schiffer (1982), Richard (1990, ...). f one isn’t trying to provide a theory of truth that can serve as a plausible theory of meaning, one can perhaps maintain that the verb-phrases in (43) and (44) are both equally true of Sam—but that (43) is more “pragmatically appropriate,” in some broad sense that covers cases that are quite unlike Grice’s (196x) animating examples. CITATIONS...Soames, Braun, Saul. And perhaps one can argue that a theory of truth should still connect ‘that’-clauses with abstracta that are individuated more finely than truth values, as opposed to extending the relevant notion of pragmatics to cover even more cases. But as we’ll see, absent reason for thinking that some truth theory is plausible as a theory of meaning that accommodates sentences like (41-48), advocates of “direct reference” versions of (D1) owe reasons for thinking that sentences of a human language (as opposed to certain expressible thoughts) are truth-evaluable. But my suspicion is that many of the debates in this vicinity are not really debates about human languages.
“hypothesis” that (43) is true if and only if Sam glonked the following ordered triple: 
⟨‘Hesperus’, ‘is’, ‘Phosphorus’⟩; where each word is taken to be a meaning-pronunciation pair. Relativizing this triple to a context doesn’t explain what it is for a long dead monolingual speaker of Sumerian to glonk an entity whose components include expressions of English. But it’s also no help to replace overtly linguistic entities with sets of logically possible worlds.

Even if we grant—pace Kripke and others—that the domain of a truth theory can include things that deserve to be called worlds in which Hesperus is not Phosphorus, we need to know what it is to deny or endorse a set of worlds-in-this-sense. (Again, the first suggestions that come to mind are not only quasi-behavioristic, they lead to contradiction given plausible assumptions.) One can say that a certain abstract entity is the “object” of a certain denial/endorsement relative to a context C just in case that entity is usable in C as a “measure” of the “content” of the denial/endorsement. But this doesn’t show that (D1) is plausible despite examples like (42-50). It shows that such examples can be accommodated by stipulation: given that the relevant verbs and complementizer phrases are true of the right things, (D1) is sustainable.

Worse, focusing on the issue of how to accommodate opacity distracts attention from the potentially massive gap between the context-invariant meaning of a sentence like (43) and the content of an assertion make by using such a sentence in a particular context. Everyone realizes that sentences like (51) show that instances of the Tarskian schema (T) are not true

(51) I denied it.

(T) True(‘S’) = p

if ‘p’ is a sentence of a tenseless metalanguage that lacks indexical/demonstrative expressions, and ‘S’ ranges over sentences of a human language. Even a sentence like (52) is tensed.

(52) There are infinitely many prime numbers.

One can say—perhaps falsely, but still comprehensively—that while there are infinitely many prime numbers today, all but finitely many of them will cease to exist when the last thinker dies. So no “eternal” sentence is truth-conditionally equivalent to (52).

As reviewed in chapter four, (D1) is not disconfirmed by the mere fact that human languages allow for tense and indices. Given work by Kaplan and others, one can describe the “character” of many context-sensitive expressions in truth-theoretic terms. But we wouldn’t be impressed by the following proposal: ‘I’, ‘it’, and the tense morpheme in (51) denote entities e, e', and e'' such that (51) is true relative to context c just in case these entities are (respectively) usable as “measures” in c of the subject, object, and time of some denial. This would trivialize the claim (D1) is compatible with the context-sensitive aspects of human languages. So the not merely technical challenge presented by sentences like (43)

(43) Sam denied that Hesperus is Phosphorus.

is to specify their alleged truth conditions in a way that meets two related constraints: we can evaluate the specifications for truth/falsity, and so make judgments about whether the relevant analogs of (T) are plausible, given our judgments about whether contextualized uses of the object language sentences are correct or not; and the dependence on context is encoded in a way that lets us see how the proposal apportions responsibility for the alleged determination of truth value, in a given context c, to the context-invariant sentence meaning and the contribution of c.

That said, I would be willing to classify (41-50) as special cases for future generations to deal with if (D1) was plausible for sentences that are in no intuitive sense “semantic”—i.e., sentences that include no lexical items like ‘true’ or ‘know’, and no phrases whose meanings are plausibly specified in terms of meanings. But in my view, (41-50) are just vivid illustrations of the far more general point that sentences of a human l-language do not have truth conditions.
In chapter three, the focus will be on action reports like (53) and (54).

(53) Al chased Theo gleefully and athletically but not skillfully.
(54) Theo chased Al gleefully and unathletically but skillfully.

As already noted, Davidson (1967a) offered an insightful suggestion about how to specify the meanings of such sentences; see §1.x above. But examples of adverbial modification still present deep difficulties for (D1). We can easily imagine scenarios in which both (53) and (54) are used to report some of what happened in a certain region of spacetime. In some of these scenarios, Al and Theo chased each other; and as we’ll see, this leads to puzzles if we assume that lexical items like ‘chase’ are true of things that are individuated as they need to be if sentences (53) and (54) are to have plausible truth conditions that are compositionally determined. For examples of this sort, involving two or more agents, there are several potential Davidsonian responses; though none are attractive. A similar point applies to other clusters of examples like (55-57).

(55) The red ball struck the green ball from the west.
(56) The green ball struck the red ball from the east.
(57) Two balls collided.

And in my view, no Davidsonian response is plausible in light of any moderately varied diet of examples, even bracketing semantic adverbs like ‘intentionally’.

One moral of these examples is that what a speaker “talks about,” by using a verb like ‘chased’ or ‘struck’, is itself dependent on the conversational context. This makes it seem even less likely that verbs like ‘denied’ are true of things in a way that is congenial to (D1). And the general point is not confined to verbs that might be described as applying to events that have multiple participants. Sentences (58-60) raise questions about what, if anything, ‘blue’ is true of.

(58) The paint is blue.
(59) The can is blue.
(60) The sky is blue.

As discussed in chapter five, the meaning of a lexical item like ‘blue’ must be flexible enough to accommodate both mass nouns like ‘paint’, count nouns like ‘can’, and whatever ‘sky’ is.

Likewise, (60) raises the questions about what, if anything, ‘sky’ is true of. The appearance of a blue sky is a complex phenomenon. Though in part because of the complexity, there is no need to posit any sky that can be a value of the variable in SKY(X) or BLUE(X). We might be able to invent a language in which BENEATH-A-BLUE-SKY(X) abbreviates a complicated scientific predicate that applies to certain regions of spacetime. But human languages have lexical items like ‘sky’ and ‘blue’, which mean what they do. I don’t think ‘sky’ is true of skies (or of sky). And I don’t think ‘chase’ is true of chases. I don’t deny that there are many chases, and that in this sense, chases exist even if skies don’t. But the existence of chases doesn’t show that ‘chase’ is true of them. I also agree that there is a sense in which there are blue skies, but no blue dragons. But it doesn’t follow that ‘sky’ is true of some things, at least not in the sense of ‘true’ that matters for a truth theory. When theorizing, there is no call to quantify over skies.

Semantic ascent does not provide a clearer view of reality. Commonsense talk does not justifiy positing the ontology required to make (D1) plausible. And while there may well be chases, over which we can quantify, there may not be enough chases to make (D1) plausible given sentences like (53) and (54). From this perspective, (60) is not yet another special case to be set aside; it highlights an issue that can be raised, perhaps less vividly, with (53-59).

Likewise, I think it is a mistake to set examples like (61) and (62)

(61) Dragons differ from unicorns.
(62) Dragons are mythical. But here is a picture of dragons chasing unicorns.
aside as marginal cases that somehow involve fiction. The idea that story-telling is a marginal use of language—while mathematized science provides a better model for ordinary talk—is obviously absurd. Given the ease with which young children engage with pretense, I think (D1) faces real difficulties here. One can insist that (61) is either false unless it contains a covert ‘in the story’ operator or true because ‘dragon’ and ‘unicorn’ can be used (like ‘Hamlet’ and ‘Polonius’) to talk about things created by acts of story-telling. But then one owes some account of sentences like (63), which raises the same issues as (42-50).

(63) In his story, Hesperus is not Phosphorus, and there is phlogiston on Vulcan.

One might want to set (61) aside, as an example of discourse that does not “aim at truth” in the relevant sense. I think this is stacking the deck in favor (D1). But the point is not just that human languages let us introduce words like ‘unicorn’ along with ‘horse’. We also have words like ‘sky’, ‘rainbow’, and ‘sunrise’. Bracketing all such expressions under the heading of ‘fiction’ overuses the heading. Moreover, an analogous point applies to proper names.

Chomsky (2000b) used the example of ‘London’, which can be used to talk about a particular location near the river Thames or a political institution that could in principle be moved. But no location can be moved, and no political institution is a location. Chomsky’s conclusion, which I accept, is that no entity is such that ‘London’ denotes (or is true of) it. Rather, there are various things we can talk about by using ‘London’ as a device for getting human beings to think about those things.

Initially, the idea of moving a city away from a river might seem recherché to readers not already familiar with the idea of moving people to Twin Earth. So let me note than in an essay on global warming in the New York Times, James Atlas—a more than competent speaker of English—discussed the “good chance that New York City will sink beneath the sea” (Nov 25, 2012). But as he noted, expecting ordinary readers to understand his words,

...the city could move to another island, the way Torcello was moved to Venice, stone by stone, after the lagoon turned into a swamp and its citizens succumbed to a plague of malaria. The city managed to survive, if not where it had begun.

And of course, there is a real chance that Venice will need to be moved as the lagoon rises. But while the argument displayed in (64) seems fine, the argument displayed in (64+) is a joke.

(64) Torcello was moved to Venice.
   Venice was a better location.
   Torcello was moved to a better location.

(64+) Torcello was moved to Venice.
   Venice was a better location.
   Venice may need to be moved.
   Torcello was moved to a better location that may need to be moved

So prima facie, there is no entity that ‘Venice’ denotes. In this respect, ‘Venice’ is like ‘Vulcan’, even though one can visit Venice but not Vulcan. Put another way, the city of Venice is not such that ‘Venice’ denotes it. However, as discussed in chapters one and six, the polysemous word ‘city’ may be linked to both a concept that applies to certain locations and a concept that applies to political institutions; and ‘Venice’ may be an instruction that can be executed by accessing a location-concept or an institution-concept. Likewise, (65)

(65) France is hexagonal, and France is a republic.
can be used to make a correct claim in many contexts. But in many such contexts, neither (66) nor (67) can be used to make a correct claim.

(66) France is a hexagonal republic.
(67) Something is both hexagonal and a republic.

So while France has many wonderful qualities, it seems that ‘France’ is not true of it.\(^{28}\)

In which case, (68) is not a marginal example. It highlights issues raised by (65) and (69).

(68) Pegasus flew across a blue sky.
(69) Secretariat ran across a dirt track

Likewise, (43) raises issues—about context sensitivity and verb meanings—

(43) Sam denied that Hesperus is Phosphorus.

that are not specific to reports of what happened in certain conversations or episodes of thinking. I think these sentences are alike in having meanings but not truth conditions. And while (41)

(41) The forty-first numbered example in this chapter is not true.

vividly highlights puzzles that attend the assumption that sentences of a human language are truth-evaluable, the real problems for (D1) are not due to ‘true’ as opposed to ‘blue’.

3.3 Theories of Truth as Theories of Meaning?

For reasons I’ll return to, challenges to (D1) are not independent of concerns about (D2).

(D1) for each human language \(H\), there is a correct theory of truth for \(H\).

(D2) a suitably formulated theory of truth for a human language \(H\)

can be the core of a correct theory of meaning for \(H\)

Though for the moment, let’s focus on (D2). For these purposes, I take it that a theory of meaning is supposed to be a theory understanding for a human language, and not merely a theory of the interpretations that the language allegedly connects with pronunciations; see Dummett (197x). But one can hypothesize that the task of providing a theory of meaning/understanding is, at its core, the task of providing a theory of truth for a human language.

If only because human languages generate interrogative sentences like (70),

(70) Bert yells?

whose pronunciation differs from that of (71),

(71) Bert yells.

it seems clear that no truth theory can itself be an adequate theory of meaning for such a language, given how meaning is related to understanding; see Dummett (197x). Even if a theory that assigns a truth condition to both (70) and (71) is correct as far as it goes, a theory of meaning must somehow encode the fact that (70) and (72)

(72) Does Bert yell?

are questions that can answered with (71) or with ‘yes’. Analogous remarks apply to (73-75)

(73) Bert yelled at midnight?
(74) Bert yelled at midnight.
(75) When did Bert yell?

Accounting for meanings of sentential “moods,” and how they are related to the meanings of complementizer phrases as in (76), is a task unto itself.

(76) I wonder whether/when/why Bert yelled.

\(^{28}\) Hornstein (1986) and Chomsky (2000b), among others, have used ‘the average man’ to show that the point is not restricted to proper nouns. This example invites responses that focus on the word ‘average’; see, e.g., Stanley and Kennedy (201x). I think the extant responses are inadequate; see, e.g., Collins (forthcoming). But even if advocates of (D) provide an unrefuted account of ‘the average man’, that doesn’t address the more general objections.
My suspicion is that truth-theoretic conceptions of meaning make the task harder than it needs to be, and that various empirical details favor the idea that meanings are instructions for how to build concepts; see Lohndal and Pietroski (2011). But I won’t press this point here. If thesis (D) (D) for each human language $H$, there is a theory of truth that is the core of a correct theory of meaning for $H$ were plausible for the “declarative fragment” of English, that would be considerable motivation for thinking that (D2) is plausible. And if (D1) turns out be quite implausible for the declarative fragment of English, then non-declarative sentences may add relatively little to the overall case against (D), even if such sentences present independent significant challenges.

Moreover, my aim is argue against (D) in ways that highlight the need for a more internalist conception of meaning according to which meanings are biologically implemented instructions for how to build concepts of a special sort. For these purposes, it is more important to focus on an objection to (D2) that Foster (197x) stressed. This challenge to (D/D2) can be raised with regard to declarative sentences, though it extends to all human linguistic expressions.

The concern is simple: non-synonymous sentences can be equivalent so far as a truth theory is concerned. Suppose that (71) and (77) are both true, at least relative to a given context.

(71) Bert yells.
(77) Ernie snores.

Given a truth theory whose theorems include (71-T) and (77-T), ignoring context for simplicity,

(71-T) $\text{True}(\text{‘Bert yells.’}) \equiv \text{Yells(Bert)}$
(77-T) $\text{True}(\text{‘Ernie snores.’}) \equiv \text{Snores(Ernie)}$

it is easy to construct another equally true truth theory that has (71-T*) and (77-T*) as theorems.

(71-T*) $\text{True}(\text{‘Bert yells.’}) \equiv \text{Snores(Ernie)}$
(77-T*) $\text{True}(\text{‘Ernie snores.’}) \equiv \text{Yells(Bert)}$

As discussed in chapter four, here are many ways of effectively supplementing a truth theory with a particular truth like (78) or (79).

(78) $\text{Yells(Bert)} \equiv \text{Snores(Ernie)}$
(79) There are finitely many prime numbers if and only if five precedes two.

And given any truth theory that has (71-T) as a theorem, it is easy to construct another equally true truth theory that has boundlessly many instances of (80) as theorems;

(80) $\text{True}(\text{‘Bert yells.’}) \equiv \text{Yells(Bert)} \& \neg$

where $\neg$ is a theorem of the truth theory. Such theorems might include (78) or (79). So prima facie, being a theorem of a truth theory doesn’t ensure—or even suggest—that the right side of the biconditional is a good specification of what the relevant object language sentence means.

This doesn’t show that thesis (D2) is false,

(D2) a suitably formulated theory of truth for a human language $H$

is the core of a correct theory of meaning for $H$
since one can hypothesize that (theorems of) a suitably formulated theory of truth will connect object language sentences with good meaning-specifiers. But as we’ll see in chapter four, there is no reason for thinking that a truth theory for a human language can be both suitably formulated and plausible. Indeed, Liar Sentences like (41)

(41) The forty-first numbered example in this chapter is not true.

allow for arguments approximating proofs that no truth theory for English can be suitably formulated and true. But the general point does not depend on there being “semanticky” lexical items like ‘true’. The point will be that even if (D2) is defensible in principle, formulating a plausible truth theory for a human language requires technical sophistication that is at odds with
the technical sophistication required to formulate a plausible meaning theory for a human language. Thesis (D) not only implies that a theory of truth (of the right sort) can serve as a theory meaning, it also implies that a theory of meaning (of the right sort) can serve as a theory of truth. And these implications remain about as implausible as they seemed to be fifty years ago.

Relatedly, while it is often useful to consider (D1) and (D2) separately when assessing truth conditional semantics, it may not be possible to evaluate these two theses independently. Given a human language H, there may not be any standard for being a correct theory of truth for H apart from the following: being extensionally equivalent to a truth theory that is the core of correct theory of meaning for H. So it may not be possible to defend (D1) and then defend the conditional claim (D2) if (D1), and then defend the conditional claim (D) if (D2) and (D1). A version of (D1) that is defensible in principle might not be viable as a theory of meaning.

It can be tempting to think that theorists can assess theorems of the form (81)

\[(81) \text{for each context } c, \text{ True}(S, c) = F(c)\]

for truth or falsity, given the judgments of native speakers, and that a theory of truth for H is correct if it is “materially adequate” in the sense of having only true theorems and having a theorem of the form (81) for each sentence S of H. But while speakers can often judge uses of sentences as correct or not, such judgments often reflect the “pragmatics” of ordinary speech—the point that a speaker was trying to convey, and not just “what was said” in a more literal sense.

More importantly, even controlling for familiar kinds of pragmatic implicatures (e.g., using ‘good’ to convey that something was not excellent), judgments regarding a particular use of a sentence S may reflect the acceptability of using S to convey a thought that is correct in the context of use. To borrow an example from Travis (200x), the truth/falsity of an utterance of (82)

\[(82) \text{The kettle is black.}\]

can depend on various aspects of the conversational situation. A given kettle might count as black on some occasions of using (82) but not others; and an entity might count as a kettle on some occasions but not others. Using ‘kettle’ and ‘black’ as we do, we can safely assume that there are many kettles and many black things. But that does ensure or suggest a general specification of the conditions in which a kettle counts as black, even if we idealize away from a familiar kind of vagueness with regard to color.\(^{29}\)

In the context of any particular discussion, perhaps one can say, ‘A kettle k counts as black if and only if k is black’. But whatever one thinks about how ‘black’ is being used in our current somewhat theoretical context, the following generalization seems wrong:

If a speaker used (82) at time t to make an assertion, in part by using ‘the kettle’ to refer to a kettle k, then that use of (82) was true if and only if k was black at t. My use of ‘black’, in formulating the generalization above, is a use in a particular context—this one—and so presumably, not a use that applies to each thing that was ever correctly described by

\(^{29}\) It is sometimes useful to think of the meaning of ‘black’ as mapping contextually determined “precisifiers” of some kind onto determinate regions of a color chart. This simplification is useful when we don’t want a discussion of semantic composition, or lexical decomposition, to get bogged down by paradoxes. But the simplification may not isolate an aspect of what ‘black’ means, as opposed to letting us ignore some complicated aspects of how ‘black’ is used. Moreover, whatever we say about small differences in color, there are further questions about how much rust, chipped paint, ornamentation (etc.) a black kettle can have. Here, the relevant notions of precisification are far less clear. And as (82) reminds us, predicates may indicate multi-dimensional notions. But if the vagueness of ‘black’ is multidimensional, with the details depending on the relevant entity—kettle or color chip—then idealizing away from vagueness is to idealize away from a host of concerns about (D1). Other color terms, apart from ‘white’, may be still more complicated.
using ‘black’ in any context. Perhaps a kettle \( k \) counts as black, on an occasion of using ‘black’, if and only if \( k \) is black relative to the standards of that occasion. But that’s not much help if the goal is to specify the meaning of ‘black’ by specifying (in a good Foster-responsive way) what ‘black’ is true of. One can say that (82) is true, relative to context \( c \), if and only if exactly one kettle is relevant in \( c \), and it is black relative to the standards of \( c \). But to evaluate this “hypothesis,” we need to know what it is for the relevant kettle to be \textit{black relative to the standards of a context}; cp. talking of denying a sentence as relativized to a context.

With such considerations in mind, Travis notes that a use of (83)

\[(D)\] There is milk in the fridge.

can be a true assertion if made in a situation that includes some spilled milk on a shelf that was supposed to be cleaned—even if other assertions of (83) do not count as true in fridge-identical situations. Suppose you want some milk for your coffee, and I sincerely utter (83), not knowing that the last of the milk was thrown out hours ago. My utterance is not true if the only milk in the fridge is a tablespoon of post-spill sour milk on the shelf; though that much fresh milk in a container might do, at least if you didn’t also want cereal. Upon reflection, it seems implausible that the meaning of (83) maps each occasion of use to a specifiable truth condition.

Of course, one can insist on a particular mapping, and try to explain away certain intuitions in terms of pragmatics. But my claim is not that (83) establishes that thesis (D) is false. The point here is simply that such examples add to the grief for (D), in part because it isn’t even clear what would count as “materially adequate” theory of truth for a language with sentences like (83). And eventually, one has to ask if any sentences actually confirm (D), instead of adding to the pile of examples that must be set aside for special treatment. Even if (51)

\[(51)\] I denied it.

can be accommodated in a plausible way, bracketing difficulties regarding ‘denied’, why think that (D) is true if (82) and (83) already present difficulties?

Bach (1994) and Recanati (2004) discuss examples like (84) and (85),

\[(84)\] The duck is ready.
\[(85)\] I prefer the rabbit.

which can be used to express different thoughts in different contexts, even holding fixed any aspects of context that are indexed by the first three words in each sentence. The duck may be ready for consumption (or consuming), for a walk, for anything that comes along, to leave, to sneeze, etc. The speaker may prefer the rabbit to the other choices for lunch, for a pet, in collars for coats, etc. One can hypothesize that all such examples are relevantly like (51), because one way or another, lexical items appear with (or just are) indices that can be used to track the relevant dimensions of context-dependence; see Stanley (2000), Rothschild and Segal (200x). I don’t think this hypothesis is plausible. But in any case, arguing that it can be maintained is not yet to argue for (D).

\[(D)\] for each human language \( H \), there is a theory of truth that is the core of a correct theory of meaning for \( H \). Proponents need to do more than provide machinery for keeping counterexamples at bay. They also need to provide reasons for thinking that (D) is true. Indeed, given (C),

\[(C)\] each human language is a biologically implementable 1-language advocates of (D) need to provide reasons for thinking that (B) is true.

\[(B)\] each human language \( H \) is a biologically implementable 1-language such that some theory of truth can serve as the core of a correct theory of meaning for \( H \). If reflection on examples made (D) plausible, then given (C), one might conclude that (B) is
plausible. But initially, (B) seems like an incredibly bold conjecture: why expect there to be theories of truth for natural (as opposed to invented) l-languages, much less truth theories that can also serve as theories of what these languages connect with pronunciations? Given sentences that confirm (D), we might embrace (B) as a wonderful discovery. But instead, every expression of a human language seems like a counterexample to (B). So why deny the appearances?

3.4 Some Motivations for Truth Theories

I have been sketching reasons for doubting (D). But for those of us who seek an alternative, it is also important to recall the motivations for it. First and perhaps foremost, when (D) was initially offered, there were few if any viable alternatives for a theory of meaning that was at least minimally compositional.

One wants to know how the meaning of a sentence S is determined by the meanings of the constituent lexical items and the way those items are arranged in S. One can fall back on the traditional idea that meanings are concepts, and say that concepts can be combined in ways reflected by sentences; see chapters one. But then one wants to know which modes of lexical composition correspond to which modes of concept composition. Moreover, it seems that it many cases, one meaning corresponds to more than one concept—both across individuals and within an individual. To take the most obvious kind of example, it seems that ‘Venus’ can be unambiguously connected with multiple concepts of Venus. Speakers may connect ‘groundhog’ with more than one concept of groundhogs. And prima facie, sets (or equivalence classes of) concepts do not compose, at least not in the ways that meanings do. So even if one is not skeptical of appeal to concepts, there was the question of how meanings can be compositional with being concepts. Davidson offered an answer inspired by Tarski; and Kripke (1963) had shown how to start applying Tarskian techniques to modal discourse.

Of course, the meaning of ‘Venus’ need not be Venus or any concept of it. The meaning might be a more abstract representation (or instruction) that is composable with others. For example, Katz & Fodor (1963) could say that individuals with different concepts still share a semantic marker. Lewis (1972) famously criticized this picture, saying “Semantics with no truth conditions is no semantics.” This slogan may have been right, given what Tarski meant by semantics.30 But it hardly follows that a theory that does not specify truth conditions is not a theory of meaning. Nonetheless, Lewis’ stipulation was taken to an important argument.

A better if less influential objection to Katz & Fodor was that they offered no account of the constraints on mappings from semantic markers to pronunciations. One wants to know why each of (1) and (2) is connected with its meaning and not the equally coherent alternative.

(1) the guest is easy to please
   (1a) It is easy for us to please the guest.
   (1b) #It is easy for the guest to please us.
(2) the guest is eager to please
   (2a) #The guest is eager that we please her.
   (2b) The guest is eager that she please us.

As Davidson noted, a truth theory is at least suggestive. One can try to specify a procedure that connects the pronunciations of (1) and (2) with the truth conditions of (1a) and (2b). Moreover, if meanings are taken to be things like concepts, one wants to know how words can combine in meaningful ways that apparently correspond to conceptual incoherence, as in the famous (86).

(86) Colorless green ideas sleep furiously.

---

30 Though see Burgess (201x) on the historical oddity of Tarski’s usage.
We know that an utterance of (86) would be bizarre, but not word salad, because we know what (86) means. Similarly, (87) has only the meaning indicated with (87a),

(87) was the hiker who lost kept waking in circles
(87a) The hiker who lost was kept walking in circles?
(87b) # The hiker who was lost kept walking in circles?
even though (87b) is a much better guess about which question a person might formulate by using an interrogative sentence that includes the words in (88) in that order.

(88) hiker lost kept walking circles

Indeed, the relevant constraint on how (87) can be understood—the auxiliary verb ‘was’ cannot be associated with the verb ‘lost’, embedded in a relative clause (see Ross 196x)—trumps considerations of coherence, as illustrated with (89).

(89) was the guest who fed waffles fed the parking meter
(89a) The guest who fed waffles was fed the parking meter?
(89b) # The guest who was fed waffles fed the parking meter?

A theory can pair the pronunciation of (89) with the right thought without offering any account of why this pronunciation is not also paired with other thoughts that have the same components.

It may not have been true that thesis (D) was the only game in town.

(D) for each human language H, there is a theory of truth that is the core of a correct theory of meaning for H

But it offered a new way of thinking about meaning. And it led, pretty quickly, to many good ideas. Following Davidson and Montague, a burst of work led to real progress in specifying the meanings of some puzzling lexical items and specific constructions, as in (90-92).

(90) Sam said that Miss Scarlet stabbed Colonel Mustard in the hall with a knife.
(91) The astronomer who hunted comets had seen every planet.
(92) Usually, water boils when the temperature rises enough.

Problems remained, even with regard to examples of “logical” vocabulary as in (93).

(93) No student will pass unless she studies.

But many good proposals about linguistic meaning were presented in truth-theoretic terms.

The techniques developed by Tarski and Church for invented languages (see chapter two) were extended to the study of tense, aspect, modality, pronouns, adverbial modifiers, causative verbs, complementizer phrases, and other aspects of human languages in new and fruitful ways. This showed that the constructional diversity of such languages, compared with the relative simplicity of any one invented language, was not itself a barrier to defending (D). Just as the mere existence of indexicals like ‘I’ did not falsify (D), likewise for the mere existence of constructions for which one could not initially think of a truth-theoretic account. With thought, apparent counterexamples suggested research strategies.

Rebutting initial objections to a proposal is not a form of confirmation. But as it became clear that examples like (90-92) could at least be accommodated, (D) came to seem like a better bet. Moreover, Lewis and Davidson initiated discussions about how facts regarding convention, communication, interpretation, and “triangulation” on aspects of a shared environment might make it the case that linguistic expressions have truth-theoretic properties, at least if human children were born with the requisite capacities and dispositions. One could imagine thinkers agreeing on a Tarski-style language, for use in communication, and imagine subsequent generations “picking it up” via some mix of nature and nurture. And even if the details remain obscure, one needn’t think that a correct “metasemantics” has been discovered to think that
extant proposals pointed in directions that were promising enough to maintain (D), at least in the absence of better ideas.

Describing meanings in truth-theoretic (or model-theoretic) terms also permitted—indeed, it was designed for—using tools from the study of logic to describe the apparent relations of implication exhibited by sentences of human languages. This makes it possible to replace the Kantian idea of one thought “containing” another with the modern notion of validity. Though in my view, this motivation is often overrated. Absent independent constraints on how the logician’s toolkit is to be used, describing analytic relations among expressions in terms of truth-preservation is unexplanatory. If S1 implies S2, there are many notions of “permissible model” such that for each such model M, S2 is “true-in-M” if S1 is. But merely providing such a notion does not explain why S1 implies S2; at best, it redescribes the explanandum in suggestive terms. Deduction is one thing, explanation another.

Put another way, just as accommodating ambiguities is unimpressive unless one also accounts for non-ambiguities, accommodating implications is unimpressive unless one also accounts for non-implications. Generating actual meaning-pronunciation pairs is easy if one is willing to tolerate overgeneration; just say that every pronunciation has every meaning. Of course, nobody thinks that every sentence implies every other sentence. But prima facie, any interesting system of logic also recognizes far more implications that ordinary speakers do. So as “Foster’s Problem” highlights, using a theory of truth as a theory of meaning requires a way saying which logical implications are not semantic implications; see chapter four.

There was also the attraction of a “spare” theory of meaning, according to which the “data” for such a theory was closely related to the target explanandum. If speakers offer judgments regarding the truth conditions of sentences, modulo context sensitivity and a little pragmatics, then it is easy to see how such judgments can be used to confirm proposed theories whose theorems specify context-sensitive truth conditions for sentences. And if such a theory can serve as the core of a meaning theory for the relevant human language, then it is corresponding easy to see how claims about linguistic meaning can be “empirically grounded.” But this thought was interwoven with an alleged motivation for (D) that is tantamount to assuming it.

Following Davidson, many advocates of (D) have taken it as given that "T-sentences" like (94) are true, modulo the context sensitivity tense.

(94) ‘Snow is white.’ is true if and only if snow is white.

But even if some invented languages permit true instances of a “disquotation” schema, it hardly follows that human languages are similar in this respect. And why should anyone who doubts (D) grant that “human T-sentences” are true?

It is often held that sentences like (94) are analytic, or at least sure to be true, because of some alleged connection between meaning and truth. In objecting to Katz and Fodor (1963), Lewis (1972) said that “we can know the Markerese translation of an English sentence without knowing the first thing about the meaning of the English sentence; namely, the conditions under which it would be true.” But why think that the first (or any) thing about a sentence meaning is “the conditions under which it would be true”? That makes for a suspiciously simple argument in favor of (D): there are boundlessly many truths like (94), including one for every sentence might initially seem problematic, including (41) and (42);

(41) The forty-first numbered example in this chapter is not true.

(42) Some ancient astronomer wondered whether Hesperus is Phosphorus and each of these (alleged) truth is an explanandum that (D) helps explain.

Indeed, if all human T-sentences are true, it is hard to see how (D1) could be false.
(D1) for each human language $H$, there is a correct theory of truth for $H$. And if every apparent counterexample is really a confirming instance of (D1), that would make (D) plausible. But it isn’t a datum that (94) is true. While there may be a sense in which disquotational T-sentences are boring, modulo context sensitivity, this doesn’t show that such sentences are true relative to contexts. Using (94) to make an assertion is weird, raising the question of what claim the speaker is making. Some ordinary speakers—as opposed to theorists who are clear about what quote marks and ‘is true if and only if’ mean in this context—may hear (94) as a variant of (95) or (96).

(95) The English sentence ‘Snow is white.’ means that snow is white.

(96) The thought that snow is white is true if and only if snow is white.

I agree that these sentences can be used to construct and express thoughts that are, at least often, correct in the relevant contexts. I also agree that (97)

(97) The English sentence ‘Snow is white.’ is true if and only if snow is white. can be used to construct and express a thought that certain theorists endorse, and that any ordinary speaker might take to be true in a context where the operating assumption is that sentences of English are true or false. But presenting speakers with sentences like (97), and noting ‘yes’ I more like response than ‘no’, shows little. If the task is viewed like one of saying ‘yes’ or ‘no’ to flash cards that contain strings like ‘$2 + 3 = 5$’ or ‘$7 \times 8 = 54$’, with (98) as a possible contrast to (97),

(98) The English sentence ‘Snow is white.’ is true if and only if snow is black. then even speakers who appreciate the difference between (97) and (95) may treat the task of evaluating (97) and (98) as that of saying ‘yes’ or ‘no’ to biconditionals like (97a) and (98b),

(97a) Snow is white if and only if snow is white.

(98a) Snow is white if and only if snow is black.

as if such speakers endorsed a “redundancy” theory of truth. In which case, disquotational T-sentences might seem like especially easy cases, or even paradigms of correct biconditionals. Still, even if some of the motivations for thesis (D) are less impressive than advertized,

(D) for each human language $H$, there is a theory of truth that is the core of a correct theory of meaning for $H$ they suggest a cluster of desiderata for any theory of linguistic meaning. In addition to wanting a theory that is compositional in some sense—the sense to be discovered—we don’t want to assume that each meaning is identical with a particular concept. We also want any theory that rejects (D) to let us reconstruct the insights, regarding specific constructions, that have been presented truth-theoretic framework new terms. We also want compensations for replacing a simple conception of word-world relations with a more complicated account that adverts to concepts. And if truth-theoretic models of implications are too permissive, we want some alternative to just falling back on talk of some thoughts containing others.

4. Summary
I called this chapter ‘Zero’ because at least with regard to the study of linguistic meaning, getting started seems to require taking a step back from where others have begun. Fundamental questions can be begged at line one. It is not obvious that meanings determine truth conditions relative to contexts, or that modulo a few niceties, (99) is true if and only if snow is white.

(99) Snow is white.

It’s certainly not obvious that meaning is use, or inferential role, or fire mixed with earth. These are compact formulations of tendentious theories. Here is another: meanings are instructions for
how to build concepts of a special sort; or abbreviating and stealing from Frege, meanings are Begriffsplans. But the various slogans raise the question of what the proposals are about.

In section one of this introductory chapter, I tried to characterize meanings in a relatively neutral way that lets us use observations regarding the much studied phenomenon of “constrained homophony” to sharpen the initial characterization. The idea is to stress that human languages—the spoken or signed languages that children naturally acquire—connect meanings with pronunciations in certain ways but not others. On the assumption that human languages are I-languages in Chomsky’s sense, the initial characterization can be sharpened further.

Some theorists will not like this assumption, despite section two. But like it or not, the human languages at least include many expression-generating procedures that operate in accord with certain constraints. If there are also human E-languages whose expressions have meanings that are not concept construction instructions, then my proposal is at least incomplete. If human I-languages generate expressions whose meanings are not Begriffsplans, then my proposal is wrong. But in evaluating any claim about what meanings are, we should ask if the claim concerns human I-languages, or an ideal language (or perhaps how we ought to use human I-languages when doing science).

Section three was an initial sketch of the currently dominant conception of linguistic meaning and some reasons for seeking an alternative. One can hypothesize that a human language connects the pronunciations of its declarative sentences with mappings from contexts to truth conditions, as specified by some theory of truth that is also a theory of meaning for the language. This bold conjecture has attractions; but it also seems to be false. The main argument in its favor of truth condition semantics is the absence of a better proposal about what meanings are. Chapter one provides an initial sketch of an alternative.
Chapter One: Introducing Concepts

In the introductory chapter, I tried to describe meanings and human languages in a way that would make the topic clear. I also advertised a few aspects of my positive proposal, according to which lexical meanings are instructions for how to access concepts of a special sort, and phrasal meanings are instructions for how to build concepts that are monadic and conjunctive. But to explain this proposal, I need to say what concepts are, and then characterize the relevant notions of monadicity and conjunction. The notion of a mental predicate will play a key role here.

1. Representation and Constituency

I assume that humans and many other animals enjoy mental representations with two striking properties: they can be used to think about things in certain ways; and they exhibit significant relations of constituency, so that a representation used to think about things in one way (e.g., as cows) can be a part of another representation used to think about things in a distinct yet analytically related way (e.g., as brown cows). I use ‘concept’ to talk about these mental entities, which can be viewed as expressions generated by languages of thought. In my view, our only remotely adequate models of concepts treat them as symbols that have adicities—valences of some sort, on analogy with hooks and eyes, or perhaps interlocking puzzle pieces—by virtue of which the symbols can combine as they do; see, e.g., Frege (1892a), Fodor (1998, 2002), Laurence and Margolis (1999), Rey (200x), Gallistel and Gibbon (2002). In this section and the next, I say that more slowly, stressing the intentionality and composability of concepts.

1.1 Thinking about Things

Let’s not argue about nomenclature. One can use ‘concept’ more permissively than I do, perhaps to include images or other representations that are not composable constituents of thoughts. One can also use ‘concept’ less permissively, perhaps to exclude representations that fail to meet certain normative constraints. Or might reserve the term for certain representable “contents,” as opposed to physically instantiated symbols. But I want to talk about a kind of physically instantiated composition that is important for cognitive science, along with a correspondingly demanding but non-normative notion of constituent, without denying that contents and nonconceptual representations are also important. In the end, one can replace ‘concept’ with a technical term characterized by its role in my account of how meaning is related to cognition. Until then, I use ‘concept’ in a restricted way; cp. Fodor (2008). Still, some readers may be unfamiliar or uncomfortable with any talk of using representations to think about things in certain ways. So some clarification, regarding aboutness and ways, is in order.

For present purposes, the relevant notion of thinking about is intentional. We can think about unicorns, even though there are no unicorns to think about. Indeed, one can posit unicorns and spend a lot of time thinking about them, constructing theories of them, diagnosing various

1 A camera may contain images of a fish, and a single image can be the result of stacking shots. Given disparate images, one can also form a less detailed “abstraction” that preserves certain features shared by inputs. But as Fodor (1998, 2003) notes, it is hard to see how any image can be used to think about fish in general, or to think about things as fish; see also Cummins (1989). An image can be associated with a concept. But a typical pet image, whatever that would be, cannot combine with a typical fish image to form a third image with which one could think about pet fish. Combining a pet image with a fish image will typically yield a mess. So it’s not plausible that images are concepts in the sense described here; cp. Prinz (2002). Likewise for prototypes; see Fodor and Lepore (2002).

2 Or at least no real unicorns. In a toy store, on in a picture book, the unicorns might be next to the teddy bears. Even if every possible world is devoid of real unicorns, (see Kripke 1980), there are actual simulacra. And in constructing certain things—toys, pictures, arguments, etc.—one may well think about unicorns. Perhaps one needs words to think about unicorns. But I see no reason for thinking that every “empty concept” reflects some word.
observations as symptoms of unicorns, etc. Similarly, one might hypothesize that a certain planet, Vulcan, passes between Mercury and the sun, and then think about this alleged planet—to estimate its mass, or wonder whether it is habitable—much as one might hypothesize that Uranus passes between a certain planet, Neptune, and the sun. A mental episode can be one of thinking about something, even if nothing is such that the episode is one of thinking about it.

One can think about countable spatiotemporal things. But one can also think about stuff, like gold or phlogiston. And the things/stuff we think about include abstracta, like numbers or intelligence. Though whenever one thinks about some stuff, or one or more things, one thinks about it or them in some way; cp. Evans’ (1982) discussion of Frege (1892b). And things/stuff can be thought about in various ways. Thinking about Hesperus can differ from thinking about Phosphorus, even though Hesperus is Phosphorus. For while the planet Venus is always itself, thinking about it as the evening star differs from thinking about it as the morning star. Thinking about woodchucks can differ from thinking about groundhogs, despite woodchucks being groundhogs; likewise for cilantro and coriander, eggplant and aubergine. Thinking about the aces can differ from thinking about the marked cards, even given a rigged deck. I take concepts to be mental representations that can be used to think about things in this intentional sense that allows for thoughts of unicorns and for Hesperus-thoughts that are not Phosphorus-thoughts.

We can still say that many concepts apply to mind-independent entities, like animals and planets, where the notion of application is not intentional. Consider an ancient astronomer who used exactly one concept C to think about Hesperus, in the intentional sense of thinking about the evening star as such, and a distinct concept C* to think about Phosphorus as such. We can say that both concepts applied to Venus: even if the astronomer would have denied it, both of his concepts applied to Hesperus, and both applied to Phosphorus. Likewise, a concept used to think about groundhogs (as such) can apply to woodchucks, which are such that any concept used to think about them applies to groundhogs. Distinct concepts, which let thinkers think about things in different ways, may apply to the same things. But thinkers can also have “empty” concepts that do not apply to anything; cp. Husserl (189x), Dummett (1981, 1994), Rey (201x). And as we’ll see, such concepts can be enormously valuable.

Let me stress that these familiar points concern thought, not how we talk about it. I grant that ‘think about’ can also be used to talk about a relation that thinkers can bear to entities/stuff to which concepts can apply; see, e.g., Burge (197x, 2010). In this “de re” sense, one can think about bosons and dark matter only if the world includes bosons and dark matter, about which one can think. Any episode of thinking de re about Hesperus is an episode of thinking de re about Phosphorus. This extensional/externalistic notion has utility, in ordinary conversations and perhaps in cognitive science, when the task is to describe representations of a shared environment for an audience who may represent that environment differently. But however this notion is related to words like ‘think’ and ‘about’, many animals have concepts that let them think about things in ways that are individuated intentionally. We animals may also have representations that are more heavily anchored in reality; see, e.g., Pylyshyn (2008). But a lot of thought is intentional, however we describe it. ³

³ If any stuff we think about is a collection of things, any episode of thinking de re about some stuff is an episode of thinking de re about a collection of countable things. But prima facie, thinking about mud (or space) as such differs from thinking about any collection of countables; see chapter five. One can be professionally suspicious of the intentional; see, e.g., Quine (1960). But prima facie, human and many other animals can think about a certain card in one way when it is face up, and another way when the same card is face down. And I don’t see how skepticism about this can be justified by reflection on how humans use phrases like ‘thinks the ace is on the left’. ³
In chapters four and six, I briefly discuss phrases like ‘think about Venus’, ‘concept of groundhogs’, and ‘doubts that Hesperus is Phosphorus’. Intentional idioms, not to be confused with intensions (i.e., procedures) or intensional logics, tell us something about meaning. But studying ordinary thought reports and concept descriptions may reveal little about the mental representations described, much as studying ‘It rained’ may reveal little that interests meteorologists. There are interesting puzzles in the vicinity; see Kripke (1979a). One wants to know how two concepts can be used to think about something in different ways: if C and C* both represent Venus, as opposed to any way of thinking about Venus, why don’t both represent Venus as such? But we face these questions about thought, whatever we say about language. And distinguishing contents from meanings may help demystify both. If lexical meanings are instructions for how to access concepts from lexical addresses, then two words can differ in meaning even if they are (or come to be) linked to a single concept/content.

In any case, however we talk about thought, we can think about things. I have said that concepts are mental symbols that can be used to think about things. So let me turn to the other key feature of concepts—viz., they exhibit significant relations of constituency.

1.2 Productive Atoms
Following custom, let’s indicate concepts with small capitals. And let’s say that COW can be used to think about cows as such. A thinker can use the concept COW to think about cows as cows.

There may well be more than one such concept, depending on what is required for thinking about things as cows.4 We can also think about cow-stuff (some of which we call ‘beef’), herds of cows, the property of being a cow, etc. So perhaps we should say that for each of many indices n, COWn is a cow-concept. But typological distinctions will soon be needed in any case; see section two. For now, let’s pretend that COW is the only concept used to think about cows as such—i.e., the only concept type whose spatiotemporally instances are so used. Likewise, let’s pretend that VENUS is the concept used to think about Venus as such, without yet encoding the idea that VENUS is about a particular thing, while COW is a more general concept. Let’s also bracket polysemy, as if BOOK is the concept used to think about books as such, even though books can be shelved or written. Before turning to these and other complications, I want to stress the crucial point that many concepts are somehow built up from others.

Thinkers who can think about brown things as such have a concept BROWN. Those who can think about brown cows as such have a concept BROWN-COW. This notation leaves room for the hypothesis that both concepts, along with COW, are conceptual atoms with no conceptual parts. (Treat the hyphen as a character, like ‘o’, that can appear in concept-designators.) The notation also allows for the hypothesis that all three concepts are composite. Perhaps COW has BEEF as a constituent, or vice versa, or the concept we use to think about brown things shares a part with the concept we use to think about brown stuff. One can speculate that COW is complex in a more substantive way, with a constituent like ANIMAL; though if COW has any such analysis, the details remain obscure.5

4 Can children think about cows this way on the basis of encountering just a few cows? How about cattle dogs, or cattle? Under what conditions would encountering a few pictures of cows suffice? As discussed below, my claims about meaning can be evaluated without taking a stand on these issues, which provide ample opportunity for terminological debate about which pairs of individuals share a given concept/word/language.

5 And pace Kantian enthusiasm for genera/differentia, GOLD cannot be decomposed into METAL and YELLOW, even if YELLOW are COLOR are analytically related. Like Fodor and Lepore (2002), I think many putative decompositions of word-level concepts are implausible; see note X. Though if formally atomic concepts can be introduced in the course of lexicalization, then some analyses may be sustainable; cp. Pietroski (2003b). In chapter two, I return to this point and its relation to Frege’s (1884) notion of analysis, illustrated with his discussion of numbers.
By contrast, brown-cow does seem to be built up from brown and cow, and not just because ‘brown cow’ is built up from ‘brown’ and ‘cow’. Prima facie, one cannot think about things as brown cows without thinking about them as cows. A thinker who lacks cow might have a concept that applies to all and only the brown cows. But it seems that no such concept could be used to think about brown cows as such. In traditional terms, it seems analytic that brown cows are cows. Simple logic, of a sort that ordinary humans appreciate, appears to ensure that cow applies to whatever brown-cow applies to. However this intuition is described, it bolsters the idea that brown-cow has simpler concepts as parts. To a first approximation, an episode of thinking about things as brown cows is an episode of thinking about things as cows and as brown things—or perhaps, to a second approximation, as cows and brown ones. Having brown-cow requires having both cow and brown, along with some capacity to combine these two concepts conjunctively; cp. Fodor and Lepore (2002) on “reverse” compositionality; see §2.1 below.

By contrast, even if one discovers or defines a concept diff that applies to all cows and differentiates them from other animals, it won’t be plausible that cow is a conjunction of animal with diff. It won’t be plausible, even to a first approximation, that to think of things as cows is to think of them as animals to which diff applies. For one can think about things as cows without thinking about them as animals. Indeed, one can suspect (without contradiction) that cows are cleverly disguised Martian robots; see Putnam (197x). While cows are animals, and any evidence to the contrary is misleading, it remains logically/conceptually possible that cows are not animals. But it is not even logically possible that brown cows are not cows.6

If brown-cow is indeed a conjunction whose constituents include cow and brown, then these concepts are somehow combinable, and likewise for many pairs of concepts. Not all complex concepts are conjunctive, even if all phrases direct construction of conjunctive concepts; see §2.1 below. But whatever modes of composition are available to humans, we can think about cows that danced, brown cows that danced on Tuesdays, dogs that chased a cow, dogs that chased every dancing cow, cows a dog chased, etc. Humans can think about things in endlessly many ways. One might suspect, as I do, that this capacity is interwoven with our capacity to generate expressions like ‘cow that danced’ and ‘cow a dog chased’. Though however the capacities emerged in natural history, humans can form boundlessly many concepts that have analyses, even if the details are unclear.

Our atomic concepts may outnumber the lexical expressions of English, even ignoring polysemy, if only because many concepts may go unexpressed. But each human has only finitely many concepts that have no conceptual constituents. Still, we can form boundlessly concepts that have distinct but systematically related contents. Given doctor, lawyer, and need, we can add doctor-who-needs-a-lawyer and lawyer-who-needs-a-doctor. Given cow and danced, we can add cow-that-needs-a-doctor, doctor-who-danced, and so on.

Moreover, this kind of concept composition is cognitively productive in a way that goes beyond mere recursive generation of representations. A thinker might be able to generate formally new concepts ad nauseum, without increasing the number of ways it can think about things. Imagine a mind that can conjoin cow with brown, conjoin the result with cow, conjoin the result with brown, etc. To be concrete, suppose that brown-cow is the complex concept

---

6 Ambiguity complicates. Sometimes, a brown cow is a root beer float. But at least one concept we can express with ‘brown cow’ is built from concepts we can express with ‘brown’ and ‘cow’. The stuff/thing contrast suggests more complexity. Like a brown house, but unlike brown paint (or root beer), a live brown cow needn’t be brown throughout; cp. Chomsky (2000b). Concepts like chicken-egg present further twists, given that at least one chicken emerged from an egg not laid by a chicken; though perhaps a chicken egg is both an egg and chicken.
+{Cow, Brown}; where ‘+’ stands for some form of conjunction, and brackets indicate a list of conjuncts. If we focus on concept types, and adopt a set-theoretic perspective, {Cow, Brown} is identical to {Cow, Brown, Cow, Brown}. But a list can include more than one token of a concept. By distinguishing first tokens from seconds, we can distinguish {‘Cow, ‘Brown} from {‘Cow, ‘Brown, ”Cow, ”Brown}. Though using concepts twice, in this way, does not provide a substantively distinct way of thinking about things.

One can insist that any distinction matters, and that thinking of things as brown cows differs from thinking of things as brown cows that are brown cows. But this difference is, to borrow Frege’s (1892b) phrase, cognitively insignificant: it is trivially analytic that brown cows are brown cows that are brown cows. By contrast, thinking about doctors who need lawyers differs from thinking about lawyers who need doctors, even if the lawyers and the doctors turn out to be the same individuals; cp. Mates (195x).

I have stressed the unbounded cognitive productivity of human concepts for two reasons. First, mere conjunction of atomic concepts cannot explain it. Second, other animals may have concepts that exhibit only more restricted kinds of productivity; see Spelke (200x), Carruthers (200x). But whatever we say about other animals, humans enjoy mode(s) of conceptual composition that yield unbounded cognitive productivity given suitable atomic concepts. This raises the question of which composition mode(s) we employ to generate phrases like ‘doctor who needs a lawyer’ and the corresponding concepts. I think we use relatively simple capacities to conjoin concepts, along with a few restricted tricks for creating mental predicates, which I take to be concepts of a special sort. But thus far, I have spoken of concepts without categorizing them. And not only are there distinctions to make, the notion of a mental predicate lies at the cusp of two historically important descriptions of thoughts and their constituents.

Let’s say that atomic thoughts have no other thoughts as parts. One classical view is that atomic thoughts are mental sentences that exhibit Subject-Copula-Predicate structure. On this view, articulated by Aristotle and the medieval logicians, predicate-conjunction plays a central role in logic. As discussed in section two and again in chapter two, I think that atomic thoughts are more varied, but that the classical tradition was importantly right about a kind of thought that is intimately connected with human I-languages. Frege (1879, 1892a)—or at least Frege as reconstructed by Church (1941, 195x) and others—offered an alternative picture of thoughts as exhibiting Function-Argument structure. As discussed in section three, this led to a much better appreciation of how relational concepts are logically related. The idea is that a unary/binary/n-ary function (from entities to truth values) can be represented with monadic/dyadic/n-adic concept that can be “saturated” by the corresponding number of “saturating” concepts, each of which indicates an entity in the function’s domain; a thought can thus indicate the (truth) value of the function given the indicated n-tuple of arguments. Monadic concepts are the analogs of classical predicates, which can be viewed as special cases of “unsaturated” thought components.7

This makes it possible to explain the notion of a mental predicate via classical and Fregean accounts of how thoughts are structured and logically related. And with a tolerably clear notion of a mental predicate in hand, I can turn (in section four) to the idea that phrasal meanings are instructions for how to build mental predicates via certain operations that severely constrain what lexical meanings can be.

7 There is a technical use of ‘predicate’ that applies to any unsaturated thought component, or least any that can be saturated by an n-tuple of unsaturable concepts. But I will avoid this usage and distinguish predicates from relations.
2. An Old Typology
Let’s say that a *general* concept can be used to think about one or more things (or some stuff) in a general way, while a *denoting* concept can be used to think about something as a particular thing (or quantity of some stuff). There are also *logical* concepts, including *quantificational* concepts that can be used with general concepts to form generalizations. In chapters five and six, I discuss *plural* concepts, *count* versus *mass* concepts, and the possibility of *neutral* concepts. But my aim is not to provide a complete typology or a substantive theory of concepts. My goal is to specify the types such that lexical and phrasal meanings are instructions for how to access and assemble concepts of those types, and to offer an account of why human languages privilege those conceptual types as opposed to others. For these purposes, it is especially important that general concepts can be *classificatory/monadic* or *relational/polyadic*: the former facilitate acts of predication, in which subjects of thoughts are represented as meeting certain conditions; the latter facilitate acts of relating things in certain ways. In thinking that Venus is a planet that orbits the sun, one might use a denoting concept to think about Venus, a classificatory concept to think about planets, and a relational concept to think about one thing orbiting another.

Ordinary language provides only a rough guide as to whether concepts are classificatory or relational. (Fathers father children; and mortals can suffer mortal wounds, as if ‘mortal’ implies a relation to a death.) Yet the distinction seems clear enough, even if application to cases is hard. This leaves it open whether any species of general concepts is more basic than any other. But there is an ancient tradition of focusing initially on classification.

2.1 Sentences, Subjects, and Predicates
I assume that there are *natural* sentences, distinct from formulae that theorists invent. The natural sentences may include some pronounceable expressions of human I-languages. But for reasons that will emerge, the classical view was that grammatical structure imperfectly reflects the Subject-Copula-Predicate structure of “deeper” sentences. Many current theories of grammatical form go farther, not even employing the notions of sentence, subject, and predicate: expressions are described as phrasal projections of lexical items, which include “functional” items (e.g., complementizers, tense morphemes, and prepositions) along with “open class” vocabulary (e.g., nouns and verbs); and no lexical category is the obvious correspondent of “Sentence” in older theories. I’ll return to the idea that a human language does not generate sentences as such, even if it generates phrases that we often call sentences. But in any case, one can grant that many *thoughts* have subjects and predicates, even if pronounceable expressions do not. There can be *languages of thought* that generate sentences in accord with procedures like the following.

Sentence → Subject copula Predicate
Predicate → PLANET, BRIGHT, COW, BROWN, PHILOSOPHER
Copula → IS, WAS
Subject → Quantifier Predicate
Quantifier → EVERY, SOME, NO
Subject → Denoter
Denoter → VENUS, Socrates

Such a procedure generates mental sentences like (1-3).

(1) [EVERYQuant PLANETPred]Subj ISCop BRIGHTPred

---

8 Aristotle spoke of “affectations of the soul.” Ockham and others explicitly posited a mental language, which would have no need for Latin’s declensions; see Pannacio (2007).
Subscripts are reminders of the typology, and outermost brackets of the form [...]Sent are suppressed for simplicity. Mental symbols may connect certain contents with biochemically realized signals that are related to perception, inference, and action. Recall that languages connect interpretations of some kind with signals of some kind; meanings and pronunciations are special cases. An animal with no capacity for external communication may have thoughts of the “categorical” forms exhibited with (1-3). Humans can generate pronounceable analogs of these sentences, despite quirks of grammar—e.g., that English permits ‘is bright’ and ‘are planets’, but not ‘is planet’. Though even if we can use (4) to express (3), the meaning of (4) may be related to (4) only indirectly.

(4) Venus is bright.

More specifically, the meaning of (4) may be an instruction for how to build a sentential concept, and executing the instruction may yield a thought that is equivalent to (3) yet formally different. As discussed in chapter two, Frege showed how a string like ‘every integer is rational’ might be used to build a thought that can be viewed as a simplification of a more ideal thought. And at least in principle, (4) might be used to build a quantificational thought like (4a) or (4b);

(4a) [EVERYQuant VENUSIZERPred]Subj ISCop BRIGHTPred
(4b) [SOMEQuant BRIGHTNESS-STATEPred]Subj ISCop EXHIBITED-BY-VENUSPred

where VENUSIZER applies to whatever VENUS denotes, and states can be viewed as boring events; cp. Quine (195x), Parsons (1990). Similarly, (5) might be used to build (5a) or (5b)

(5) Every planet rose.

(5a) [EVERYQuant PLANETPred]Subj ISCop ONE-THAT-ROSEPred
(5b) [EVERYQuant PLANETPred]Subj WASCop RISER-IN-A-RISINGPred

via some procedure that underlies a distinctively human form of unbounded cognitive productivity. For now, though, it will do no harm to pretend that (4) is a simple case in which grammatical structure directly reflects conceptual structure.

From this perspective, (4) has as a subject in the derivative sense of having a constituent that indicates the subject of (3). But thoughts, or least episodes of thinking, can have subjects in a sense that is not derivative or merely formal. One can think that Venus is bright by perceiving or remembering Venus, and thereby thinking about it in a certain way, while classifying it as bright. A mental act of this sort has Venus as its subject matter, and not just because VENUS Den and BRIGHTPred differ formally. Using classical terminology, we can say that the judgment is one in which brightness is predicated of Venus, which can be thought about independently of any predication. But this might suggest that BRIGHTPred denotes brightness, and that predicates are special cases of denoters; cp. Frege (1892a), discussed in chapter two. To avoid this suggestion, it may be better to say that the judgment is predicational, with brightness as the content of the predication, and of Venus; where there can also be thoughts of Vulcan (see §1.x above).

I don’t deny that sentences like (4) can be used in mental acts, say of hypothetical reasoning, that are less directly about Venus. And a quantificational sentence may have its

---

9 At least not if ‘planet’ is used as a count noun; though a bumbling creator might say, ‘There is mud all over the floor, and there is planet all over the yard’. Many other human languages, including ancient Greek, do not require that an unplural count noun appear with an article.

10 Pace Russell (1912). I don’t think there were ever any good reasons for denying that humans and other animals have singular concepts of distal objects. But in light of Kripke (1980), Evans (1982), and many others, it seems that intuitive considerations bolster other reasons for positing such concepts. Citations to Pylyshyn and Burge.
subject in a derivative way; perhaps (1) is about the planets only because it predicates brightness of each planet and nothing else. On the other hand, if (1) restricts the domain of quantification to the planets—with brightness predicated of everything in the relevant domain—then the planets may be the subject matter of (1), just as Venus is the subject matter of (2) and (3). But even if some mental sentences have subjects in a derivative sense, any grammatical notion of subject seems to be more removed from the notion of a potential target for predication.

One can say that the grammatical subjects of (6) and (7) are pleonastic—

(6) It rained yesterday.
(7) It was easy to please the guest.

i.e., they do not indicate subjects of corresponding thoughts. One can also say that in (8),

(8) The guest was easy to please.

‘the guest’ is both the non-pleonastic subject of (8) and part of its tensed predicate. But then talk of grammatical subjects seems to be (at best) a way of saying that tensed clauses like (4-8) have a “left edge constituent” that somehow makes them suitable for expressing complete thoughts.

By contrast, “mere phrases” like (9) and (10) feel incomplete if used to express thoughts.

(9) rained yesterday
(10) easy to please the guest

Correlatively, ‘it’ seems to be logically irrelevant in (7), which is somehow equivalent to (8). And while (11) somehow implies that something barked yesterday,

(11) It barked yesterday.

(6) does not likewise imply that something rained yesterday. The phenomenon is real. But the notion of a grammatical subject, like the notion of a complete sentence, seems to depend on the idea of (completely) expressing a thought.

More generally, it has long been clear that grammatical forms do not perfectly reflect the logical forms of thoughts, whatever these forms are. The ancient Greeks were aware of sophisms like the following: that dog is a father, and that dog is yours; so that dog is your father. This inference, obviously bad, differs somehow from a superficially parallel but good inference: that dog is a mutt, and that mutt is yours; so that dog is your mutt. (See Plato, Euthydemus 298 d-e.) This raises the question of what makes good inferences good, and Aristotle’s partial answer seems fundamentally correct. Sentences exhibit valid forms of inference.

The relevant form exhibited by (12-14) can be described in various ways.

(12) \([\text{EVERY}_{\text{Quant}} \text{HUMAN}_{\text{Pred}}]_{\text{Subj}} \text{IS}_{\text{Cop}} \text{ANIMAL}_{\text{Pred}}\)
(13) \([\text{EVERY}_{\text{Quant}} \text{ANIMAL}_{\text{Pred}}]_{\text{Subj}} \text{IS}_{\text{Cop}} \text{MORTAL}_{\text{Pred}}\)
(14) \([\text{EVERY}_{\text{Quant}} \text{HUMAN}_{\text{Pred}}]_{\text{Subj}} \text{IS}_{\text{Cop}} \text{MORTAL}_{\text{Pred}}\)

Using ‘Φ’, ‘Ψ’, and ‘Ω’ as placeholders for predicates, we might say that any instance of the ordered triple (15) is such that the first two sentences jointly imply the third.

(15) \(\langle [\text{EVERY}_{\text{Quant}} \Phi]_{\text{Subj}} \text{IS}_{\text{Cop}} \Psi]_{\text{Sent}}, [\text{EVERY}_{\text{Quant}} \Psi]_{\text{Subj}} \text{IS}_{\text{Cop}} \Omega]_{\text{Sent}}\)

\(\text{Sent} \supset \text{NOT} \text{Sent} \)
\(\text{Sent} \supset \text{Sent} \text{AND} \text{Sent} \)
\(\text{Sent} \supset \text{Sent} \text{IF} \text{Sent} \)

Then one can say that instances of (16) are truisms.\(^{11}\)

\(^{11}\) This also makes it possible to formulate schematic principles of propositional logic. S1 and S2, so S2 and S1;
which raise the question of how one can refer to those sentences and truth evaluable judgments/claims, and the relevance of various c

In this sense, sentences (27) and (28) are equivalent.

S2 if S1, and not-S2, so not-S1; etc. Stoic logicians also discussed structural ambiguities, the differences between sentences and truth evaluable judgments/claims, and the relevance of various contextual factors—including deaths, which raise the question of how one can refer to those who no longer exist; see Bobzien (1998).
By exploiting such equivalences, and sometimes proving a conditional by *reductio*—showing that the antecedent cannot be consistently conjoined with the negation of the consequent—Aristotle showed that the valid categorical syllogisms can be derived from (19-22), which came to be known as Barbara, Celarent, Darii, and Ferio; see Parsons (2014) for detailed discussion.

To take a trivial example, given that no $\Omega$ is $\Psi$ if and only if no $\Psi$ is $\Omega$, (29) can be derived from (30).

$$\begin{align*}
(29) \quad \text{every } \Phi & \text{ is } \Psi \\
\text{no } \Omega & \text{ is } \Psi \\
\text{no } \Psi & \text{ is } \Omega \\
\text{no } \Phi & \text{ is } \Omega
\end{align*}$$

More interestingly, given that (23) contradicts (24), (31) can be derived by *reductio* from (19);

$$\begin{align*}
(31) \quad \text{some } \Phi & \text{ isn’t } \Psi \\
\text{every } \Omega & \text{ is } \Psi \\
\text{some } \Phi & \text{ isn’t } \Omega
\end{align*}$$

To affirm the premises of (31) while denying its conclusion is to affirm each of the following: some $\Phi$ isn’t $\Psi$; every $\Omega$ is $\Psi$; every $\Phi$ is $\Omega$. Or replacing the first element of this trio with its equivalent, and the reversing the order: every $\Phi$ is $\Omega$; every $\Omega$ is $\Psi$; not every $\Omega$ is $\Psi$. But any trio of this form would violate (19). So we need not view (31) as a basic form of valid inference.

### 2.2 Predicates, Conjunction, and Reduction

I am reviewing some classical logic, in order to provide some initial grip on the notion of predicate I have in mind when I say that meanings are instructions to build mental predicates.

Aristotle’s logic suggested that predicates figure in certain patterns of inference that humans find naturally compelling. He was especially interested in logical relations exhibited by two or more generalizations. But as he recognized, (32) is also compelling;

$$\begin{align*}
(32) \quad \delta & \text{ is } \Psi \text{ if every } \Phi \text{ is } \Psi, \text{ and } \delta = \Phi
\end{align*}$$

where ‘$\delta$’ corresponds to denoting concepts like *Venus*. This is, of course, the syllogism relevant to the inference from (1) and (2) to (3).

$$\begin{align*}
(1) & \quad [\text{EVERY}_{\text{Quant}} \text{ PLANET}_{\text{Pred}}]_{\text{Subj}} IS_{\text{Cop}} BRIGHT_{\text{Pred}} \\
(2) & \quad [\text{VENUS}_{\text{Den}}]_{\text{Subj}} IS_{\text{Cop}} \text{ PLANET}_{\text{Pred}} \\
(3) & \quad [\text{VENUS}_{\text{Den}}]_{\text{Subj}} IS_{\text{Cop}} BRIGHT_{\text{Pred}}
\end{align*}$$

There also seem to be cases in which two predicates are *joined* to form a third.

Prima facie, we can use ‘*Venus is a brown cow*’ to express an analog of (33), which implies (34).

$$\begin{align*}
(33) & \quad [\text{VENUS}_{\text{Den}}]_{\text{Subj}} IS_{\text{Cop}} [\text{BROWN}_{\text{Pred}} \text{ COW}_{\text{Pred}}]_{\text{Pred}} \\
(34) & \quad [\text{VENUS}_{\text{Den}}]_{\text{Subj}} IS_{\text{Cop}} \text{ COW}_{\text{Pred}}
\end{align*}$$

Relatedly, it seems that (35) is a truism, while (36) is not.

$$\begin{align*}
(35) & \quad [\text{EVERY}_{\text{Quant}} [\text{BROWN}_{\text{Pred}} \text{ COW}_{\text{Pred}}]_{\text{Pred}}]_{\text{Subj}} IS_{\text{Cop}} \text{ COW}_{\text{Pred}} \\
(36) & \quad [\text{EVERY}_{\text{Quant}} \text{ COW}_{\text{Pred}}]_{\text{Subj}} IS_{\text{Cop}} [\text{BROWN}_{\text{Pred}} \text{ COW}_{\text{Pred}}]_{\text{Pred}}
\end{align*}$$

As noted in section one, [[BROWN]_{\text{Pred}} \text{ COW}_{\text{Pred}}]_{\text{Pred}} seems to be a conjunction of its constituents. And as medieval logicians knew, such predicates interact with quantifiers in interesting ways.

Suppose that some brown cow is a dumb animal. Then some cow is a dumb animal, and some brown cow is an animal. Conversely, if no cow is dumb, then no brown cow is dumb, and no cow is a dumb animal. In a sentence of the categorical form (37),

$$\begin{align*}
(37) & \quad [\ldots]_{\text{Quant}} [\ldots]_{\text{Pred}}]_{\text{Subj}} IS_{\text{Cop}} \ldots \text{Pred}
\end{align*}$$

the indefinite quantifier *some* licenses *reduction* of a conjunctive predicate—e.g., replacing [[BROWN]_{\text{Pred}} \text{ COW}_{\text{Pred}}]_{\text{Pred}} with \text{COW}_{\text{Pred}}—in both the “intrasubject” position and the “copularized” position, while the negative quantifier *no* licenses *expansion* of a predicate in both positions.
Intuitively, reduction creates a logically weaker predicate, while expansion creates a logically stronger predicate: \([\text{SOME}_{\text{Quant}} \ [\text{BROWN}_{\text{Pred}} \ \text{COW}_{\text{Pred}}]_{\text{Pred}}]_{\text{Subj}} \ \text{IS}_{\text{Cop}} \ \text{DUMB}_{\text{Pred}}\) is logically stronger than \([\text{SOME}_{\text{Quant}} \ \text{COW}_{\text{Pred}}]_{\text{Subj}} \ \text{IS}_{\text{Cop}} \ \text{DUMB}_{\text{Pred}}\); but \([\text{NO}_{\text{Quant}} \ \text{COW}_{\text{Pred}}]_{\text{Subj}} \ \text{IS}_{\text{Cop}} \ \text{DUMB}_{\text{Pred}}\) is logically weaker than \([\text{NO}_{\text{Quant}} \ [\text{BROWN}_{\text{Pred}} \ \text{COW}_{\text{Pred}}]_{\text{Pred}}]_{\text{Subj}} \ \text{IS}_{\text{Cop}} \ \text{DUMB}_{\text{Pred}}\).

As the Stoics (see note 11) had observed, conjunct reduction seems to be the default direction of inference for sentences: S1 and S2; so S1. When sentences are offered without any explicit connectives, as in a story, the default expectation is that each claim is offered as an addition to those already presented. But this pattern can be reversed with (classical) disjunction: S1; so either S1 or S2. Likewise, sentential negation licenses expansion in its scope: not S1; so not both S1 and S2. This pattern apparently extends to predication. If one of the brown cows is one of the dumb animals, then one of the cows is one of the animals; but if Venus isn’t a cow, then Venus isn’t a brown cow. In one sense, this is hardly surprising. Who would want the expectation to be that lengthening a story (or a list of premises) reduces the content? But thinking about how quantifiers relate to predicate reductions/expansions, thereby focusing on minimally different pairs of sentence, turned out to be a good idea.

One can view the so-called *dicta de omni et nullo* in this light: the default (all-purpose) rule is to permit predicate reduction; though in special cases, with negation as the prototype, the direction of valid inference is inverted. We can imagine mental languages that work differently. But if it is analytic that \([\text{COW}_{\text{Pred}}]_{\text{Subj}} \ \text{IS}_{\text{Cop}} \ \text{DUMB}_{\text{Pred}}\) applies to whatever \([\text{BROWN}_{\text{Pred}} \ \text{COW}_{\text{Pred}}]_{\text{Pred}}\) applies to—while the converse assumption is substantive, and equivalent to (36)—then minds that deploy such concepts may well be disposed to treat reduction as the “normal” direction of inference. Correlatively, one might expect minds to treat junctions of predicates as conjunctions. In any case, it is interesting that in terms of implications involving complex predicates, the universal quantifier EVERY is like NO with regard to the subject position, but like SOME with regard to the copularized position: if every cow is an animal, then every brown cow is an animal; yet if every cow is a dumb animal, then every cow is an animal.

Let concatenation of predicate variables as in ‘some \(\Omega \Phi\)’ signify a complex predicate like ‘brown cow’, in which two predicates have been suitably conjoined; and ignore any effects of linear order, like the ungrammaticality (in English) of ‘some cow brown is an animal’. Then we can list six generalizations like the following: if some \(\Omega \Phi\) is \(\Psi\), then some \(\Phi\) is \(\Psi\); etc. But these generalizations can be perspicuously summarized with the table below.\(^\text{12}\)

---

\(^\text{12}\) See Ludlow’s (2002) summary and development of Sanchez (19xx), which reviews relevant medieval achievements. I am greatly indebted to Ludlow’s discussion. Within linguistics, Ladasaw (1980) animated discussions of these quantifier properties by observing that so-called “negative polarity items”—expressions like ‘ever’ and the idiom ‘give … a plug nickel’, which seem to require a negative context (as in ‘I wouldn’t give you a plug nickel for this book, and you should never give it’)—are licensed in the “inverting” environments (now described as “downwardly entailing”), but not in the default environments that license predicate reduction.

\(\text{No}/\text{Every}/*\text{Some cow that ever grazed} is an animal.\)

\(\text{No}/\text{Every}/*\text{Some cow is an animal that ever grazed.}\)

Though as Ludlow stresses, the quantifier MOST does not license reduction or restriction in the subject position. It may be that most (of the) brown cows are dumb, while most cows aren’t dumb; and most cows can be dumb, even if most brown ones aren’t. Still, ‘Most cows that ever grazed are dumb’ is fine. Ludlow offers a specific proposal according to which a quantifier can mark one of its predicate positions in the non-default way, so that reduction is not licensed, without thereby inverting the default direction of inference; cp, Ladusaw (1996).
From this perspective, Aristotle’s universal quantifier is governed by two inferential principles that characterize its distinctive logical role.

\[(38) \text{ every } \Phi \text{ is } \Omega \Psi \]
\[(39) \text{ every } \Phi \text{ is } \Psi \]

Principle (38) reflects the fact that \textit{EVERY} leaves the copularized (“external”) predicate position, corresponding to ‘\(\Psi\)’, in the default state of licensing \textit{reduction}—e.g., \([\text{BROWN}_{\text{Pred}} \text{COW}_{\text{Pred}}]_{\text{Pred}}\) to \(\text{COW}_{\text{Pred}}\). Principle, (39), reflects the fact that with regard to intra-subject (“internal”) predicate position, corresponding to ‘\(\Phi\)’, \textit{EVERY} inverts the normal inferential polarity and licenses \textit{expansion}—e.g., \(\text{COW}_{\text{Pred}}\) to \([\text{BROWN}_{\text{Pred}} \text{COW}_{\text{Pred}}]_{\text{Pred}}\). We can view (38) as an “elimination rule” that can go unstated in a system that permits reduction in the absence of any quantifier-specific rule. But in such a system, (39) is a substantive “introduction” rule, which captures the idea that instances of (40) permit the inverted direction of inference for predicates in the ‘\(\Phi\)’-position.  

\[(40) \text{ every } \Phi \text{ is } \Psi \]

I’ll return to another sense in which an instance of (40) warrants addition of its external predicate to its internal predicate in \textit{default} positions, as in (41) and the more interesting (42).

\[(41) \text{ every } \Phi \text{ is } \Psi \text{ and every } \Phi \text{ is } \Psi \]
\[(42) \text{ some } \Phi \text{ is } \Psi \text{ and every } \Psi \text{ is } \Omega \]

But first, I want to stress the relevance of (38) and (39) for reducing many valid inference patterns to a small stock of basic principles governing conjunctive predicates.

\[(38) \text{ every } \Phi \text{ is } \Omega \Psi \]
\[(39) \text{ every } \Phi \text{ is } \Psi \]
\[(40) \text{ every } \Phi \text{ is } \Psi \]

For it turns out that the predicates figure centrally in a rather elegant and arguably natural logic that philosophers of language often forget about; though see Sommers (198x), Parsons (2014).

According to Aristotle, (43) implies that there is at least one \(\Phi\);

\[(43) \text{ every } \Phi \text{ is } \Phi \]

and perhaps The Philosopher was correct. If (44) likewise implies that there is some \(\Psi\Phi\),

\[(44) \text{ every } \Psi\Phi \text{ is } \Phi \]

then (43) and (44) are not strictly equivalent. Still, instances of (44) seem truistic; recall (35).

\[(35) \{[\text{EVERY}_{\text{Quant}} \text{BROWN}_{\text{Pred}} \text{COW}_{\text{Pred}}]_{\text{Pred}}\}_{\text{Subj}} IS_{\text{cop}} \text{COW}_{\text{Pred}}\]

And in any case, (44) can be derived from (45), given (38) and (39).

\[(45) \text{ every } \Phi \text{ is } \Psi\Phi \]

I take it that (46) is even more obviously truistic;

\[(46) \text{ some } \Phi \text{ is } \Psi\Phi \text{ if some } \Phi \text{ is } \Psi \]

\[13 \text{ We can view \textit{SOME} as governed by two default elimination rules, and \textit{NO} as governed by two introduction rules, both of which need to be stated explicitly. But my suggestion is not that contents of quantificational concepts (much less the meanings of corresponding words) are fully characterized by any Gentzen-style rules of inference. For one thing, there are various indefinite quantifiers; \textit{TWO}, \textit{SEVERAL}, etc. And while the contents and inferential roles of logical concepts are clearly related, I take no stand on how the content of \textit{EVERY} is related to its effect/non-effect on the default inferential polarities of predicate positions, not even for instances of (40).} \]
and given (46), (21) can be used to derive the intuitively similar (21+).\(^{14}\)

\[
\begin{array}{ll}
(21) & \text{some } \Phi \text{ is } \Psi \\
& \text{every } \Psi \text{ is } \Omega \\
& \text{some } \Phi \text{ is } \Omega \\
\hline
(21+) & \text{some } \Phi \text{ is } \Psi \\
& \text{every } \Phi \text{ is } \Omega \\
& \text{some } \Phi \text{ is } \Omega
\end{array}
\]

This permits reduction (20) and (22), Celarent and Ferio, to (21).

\[
\begin{array}{ll}
(20) & \text{every } \Phi \text{ is } \Psi \\
& \text{no } \Psi \text{ is } \Omega \\
& \text{no } \Phi \text{ is } \Omega \\
\hline
(22) & \text{some } \Phi \text{ is } \Psi \\
& \text{no } \Psi \text{ is } \Omega \\
& \text{no } \Psi \text{ is } \Omega
\end{array}
\]

Proofs by \textit{reductio} are sketched below, with line (c) reflecting negations of the conclusions.

\[
\begin{array}{ll}
(a) & \text{every } \Phi \text{ is } \Psi \\
(b) & \text{no } \Psi \text{ is } \Omega \\
(c) & \text{some } \Phi \text{ is } \Omega \\
(d) & \text{some } \Phi \text{ is } \Phi \Phi \\
(e) & \text{some } \Omega \Phi \text{ is } \Phi \\
(f) & \text{some } \Omega \Phi \text{ is } \Psi \\
(g) & \text{some } \Psi \text{ is } \Omega \Phi \\
(h) & \text{some } \Psi \text{ is } \Omega \\
(i) & \bot
\end{array}
\]

So given the rules for reducing/adding predicates, Aristotle’s syllogisms can be derived from (19) and (21), Barbara and Darii.

\[
\begin{array}{ll}
(19) & \text{every } \Phi \text{ is } \Psi \\
& \text{every } \Psi \text{ is } \Omega \\
& \text{every } \Phi \text{ is } \Omega \\
\hline
(21) & \text{some } \Phi \text{ is } \Psi \\
& \text{every } \Psi \text{ is } \Omega \\
& \text{some } \Phi \text{ is } \Omega
\end{array}
\]

Moreover, (19) and (21) can themselves be viewed as applications of the special license provided by a universal generalization to \textit{add} its external predicate to its internal predicate in default positions. Recall (41) and (42).

\[
\begin{array}{ll}
(41) & \text{every } \Phi \text{ is } \Psi \\
& \text{every } \Phi \text{ is } \Phi \Phi \\
\hline
(42) & \text{some } \Phi \text{ is } \Psi, \text{ and every } \Psi \text{ is } \Omega \\
& \text{some } \Phi \text{ is } \Omega \Psi
\end{array}
\]

And consider the following chains of reasoning, which are intimately related to (21).

\[
\begin{array}{ll}
(a) & \text{some } \Phi \text{ is } \Psi \\
(b) & \text{every } \Psi \text{ is } \Omega \\
(c) & \text{every } \Psi \text{ is } \Omega \Psi \\
(d) & \text{some } \Phi \text{ is } \Omega \Psi \\
(e) & \text{some } \Phi \text{ is } \Omega \\
\hline
(a) & \text{some } \Phi \text{ is } \Psi \\
(b) & \text{every } \Psi \text{ is } \Omega \\
(c) & \text{every } \Psi \text{ is } \Omega \Psi \\
(d) & \text{some } \Phi \text{ is } \Omega \Psi \\
(e) & \text{some } \Phi \text{ is } \Omega
\end{array}
\]

The move from (b) to (d) on the right looks like a variant of (21) itself, and in one sense it is. But instead of viewing this step as warranted by a syllogism, we can view it as an application of a conversion licensed by (b): adding \(\Omega\) to \(\Psi\) is permissible, even in a default position where predicate expansion is not licensed in general.\(^{15}\) Similarly, the chain of reasoning below can be viewed as an elaborated—or perhaps fully spelled out—variant of (19).

\[\text{\footnotesize \(14\)} \text{ Given the first premise of (21+) and (46), it follows that } \text{some } \Phi \text{ is } \Psi \Phi. \text{ Given the second premise of (21+) and (39), it follows that } \text{every } \Psi \Phi \text{ is } \Omega. \text{ So given (21), it follows that } \text{some } \Phi \text{ is } \Omega.\]

\[\text{\footnotesize \(15\)} \text{ If it helps, think of (42) as a limited rule of “existential instantiation” that permits conversion of } \‘\Phi a & \Psi a\’ \text{ into } \‘\Phi a & \Omega a\’ \text{ given (b).}\]
(a) every $\Phi$ is $\Psi$ [premise]
(b) every $\Psi$ is $\Omega$ [premise]
(c) every $\Phi$ is $\Psi\Phi$ [a, 41]
(d) every $\Psi$ is $\Omega\Psi$ [b, 41]
(e) every $\Phi$ is $\Omega\Psi\Phi$ [c, d]
(f) every $\Phi$ is $\Omega\Psi$ [e]
(g) every $\Phi$ is $\Omega$ [f]

Here, line (c) reflects the very license that (a) provides: $\Psi$ can be added to $\Phi$, even in a default position where predicate expansion is not licensed in general. Likewise, (d) reflects the license that (b) provides: $\Omega$ can be added to $\Psi$. Line (e) simply reflects the fact that these licenses are agglomerative: if you can add $\Psi$ to $\Phi$, and you can add $\Omega$ to $\Psi$, you can do both. Not all permissions are agglomerative. Given diet or manners, it may be permissible to have cake, and permissible to have pie, but not permissible to have both. Given law and an empty gun, it may be permissible to load the gun with bullets, and permissible to be pull the trigger, but not permissible to do both (in that order). Put another way, some permissions/obligations are defeasible; and as Hory (201x) shows, there may be governed by an interesting logic. But one can view the medieval logicians as getting at a less nuanced system in which (47) doesn’t imply (48), but (48) does imply (49).

(47) $[\text{EVERY}_{\text{Quant}} \text{BIRD}_{\text{Pred}}]_{\text{Subj}} IS_{\text{Cop}} \text{WINGED}_{\text{Pred}}$
(48) $[\text{EVERY}_{\text{Quant}} \text{BIRD}_{\text{Pred}}]_{\text{Subj}} IS_{\text{Cop}} \text{ONE-THAT-FLIES}_{\text{Pred}}$
(49) $[\text{EVERY}_{\text{Quant}} \text{FIFTY-POUND}_{\text{Pred}} \text{BIRD}_{\text{Pred}}]_{\text{Pred}} IS_{\text{Cop}} \text{ONE-THAT-FLIES}_{\text{Pred}}$

This is not especially interesting as logic. But given the intuitive force of Aristotelian syllogisms, which form a network that can be reduced to small number intuitively compelling principles, it seems that predicates play a central role in an interesting form of thought that is governed by a “natural logic” that treats predicate-reduction as a privileged form of inference. At a minimum, we can draw on this ancient tradition to characterize a notion of predicate that is independent of modern logic and contemporary conceptions grammatical structure. This is what I have in mind when I say that phrasal meanings are instructions for how to build mental predicates. One can imagine animals who are able to generate categorical thoughts, in a limited way, but then acquire a capacity to form predicates in a cognitively productive way—thereby becoming able to construct boundlessly many instances of (40) that are systematically related.

(40) every $\Phi$ is $\Psi$

2.3 Predicates and Relations

It is, however, no part of my view that all thoughts are categorical. On the contrary, I assume that humans and many other animals have thoughts that involve irreducibly relational concepts. We may have many ways of thinking that one thing caused another. But in thinking that a certain flash caused a certain bang, we presumably use a relational concept of causation. Likewise, I assume that we use relational concepts in thinking that one thing is above another, that Venus is between Mercury and Earth, etc. Let’s indicate relational concepts with subscripts that reflect the relevant number of relata—the “arity” of the relation, ignoring tense for simplicity—as follows: $\text{ABOVE}_2$, $\text{BETWEEN}_3$, etc. And let’s assume that a relational concept can combine with the relevant number of singular/denoting concepts to form a complete thought. In section three, I cash this out in terms of Frege’s (1892a) notion of saturation. Though whatever the details, the point is that there are many noncategorical thoughts, which we can represent with a familiar kind of notation: $\text{ABOVE}_2(\text{VENUS}_{\text{Den}}, \text{BESSIE}_{\text{Den}})$; $\text{BETWEEN}_3(\text{VENUS}_{\text{Den}}, \text{MERCURY}_{\text{Den}}, \text{EARTH}_{\text{Den}})$, etc.
A thinker might be able to use some relational concepts, in limited ways, to form complex predicates like \textit{ONE-THAT-IS-ABOVE-BESSIE}\textsubscript{Pred} and \textit{ONE-THAT-CAUSED-A-BANG}\textsubscript{Pred}. As discussed in section three, such use is not trivial; forming “partially saturated” concepts like \textit{ABOVE}_2(\_, BESSIE\textsubscript{Den}) is no small matter. But in any case, these complex predicates are not mere conjunctions of simpler ones. They presumably have relational concepts as constituents.

It may be that some of these relational concepts are reducible to others. Perhaps some analog of \textit{CAUSED}_2(\_, \_) can be reduced to a cluster of concepts that includes \textit{BEFORE}_2 and a “second-order” concept that can combine with a pair of predicates. One can imagine a mind that uses the classical quantifier \textit{EVERY}\textsubscript{Quant} to introduce a concept \textit{EACH}_2 that can appear in thoughts like \textit{EACH}_2(COW\textsubscript{Pred}, BROWN\textsubscript{Pred}). One can also imagine a mind that uses \textit{EACH}_2 to introduce \textit{EVERY}\textsubscript{Quant}. And a sophisticated mind might be able to use a second-order concept to introduce a first-order concept \textit{HUME}_2(\_, \_) that can appear in thoughts like \textit{HUME}_2(FLASH\textsubscript{Den}, BANG\textsubscript{Den}); where the relational concept applies to pairs of things such that for some predicates \(\Phi\) and \(\Psi\), the first thing is \(\Phi\), the second thing is \(\Psi\), and every \(\Phi\) is \(\Psi\). But any such reduction of \textit{CAUSED}_2(\_, \_) requires a stock of basic relational concepts and the resources required to introduce/define others.

As discussed in chapter two, Frege offered a framework for such introduction/definition. He also showed how a certain kind of thinker could use relational concepts, along with concepts of truth values, to introduce \textit{analogs} of classical predicates. A “deeply relational” thinker, who has no categorical thoughts, can mimic a thinker who has such thoughts and reasons with them in an Aristotelian fashion. And we view classical logic as a (very) limiting case of Frege’s far more powerful logic, which was designed to accommodate higher-order relational thought and arguments involving arithmetic \textit{induction}, as opposed to mere syllogisms. But this doesn’t show, or even suggest, that humans are deeply relational thinkers.

Frege assumed that we naturally think and talk in a subject-predicate format, and that we need help—of the sort provided by his invented Begriffsschrift—in order to use our rudimentary capacities for relational thought in systematic ways. I think this is basically right: our categorical thoughts are governed by a natural logic that lets us appreciate certain implication relations among predicates; but our relational concepts are related in less systematic ways. We use relational concepts in natural modes of thought; but we do not reason with such concepts in the manner of an ideal Fregean thinker. Indeed, it seems that given relational concepts like \textit{BETWEEN}_3, we go the trouble of making predicates like \textit{BETWEEN-MERCURY-AND-EARTH}\textsubscript{Pred}. As discussed in chapter five, many facts suggest that we use human languages as tools for connecting relational concepts with formally distinct but analytically related predicates that we reason with in natural systematic ways. In my view, event analyses often reflect this.

If (6) is used to assemble and expression a thought of sort indicated with (6a),

\textit{(6) It rained yesterday.}

(6a) \textit{SOME}\textsubscript{Quant} \textit{THING}\textsubscript{Pred} \textit{WAS}\textsubscript{Cop} \textit{RAIN-EVENT}\textsubscript{Pred} \textit{YESTERDAY-EVENT}\textsubscript{Pred} \textit{Pred}

then ‘yesterday’ corresponds to a predicate that can be validly eliminated. And perhaps (50)

\textit{(50) Sadie chased Bessie.}

is used to assemble and expression a thought of sort indicated with (50a),

(50a) \textit{SADIE}\textsubscript{Den} \textit{WAS}\textsubscript{Cop} \textit{CHASER-IN-OF-Chasing-OF-BESSIE}\textsubscript{Pred}

even if the predicate in (50a) is formed by using the relational concept in (50b).

(50b) \textit{CHASED}_2(\textit{SADIE}\textsubscript{Den}, \textit{BE BESSIE}\textsubscript{Den})

I’ll have much more to say about this below. But to repeat, my view is not that all thoughts are categorical. On the contrary, I think humans use meanings to \textit{make} predicates from ingredients.
that include many relational concepts. And while I think classical logic has a natural basis that Frege transcended, I don't think classical logic provides an adequate theory of natural thought.

Towards the end of the eighteenth century, Kant could say (without much exaggeration) that logic had followed a single path since its inception, and that ‘since Aristotle it has not had to retrace a single step’. But he just wrong to say that classical logic was ‘to all appearance complete and perfect’.  

As the medieval logicians knew, relations presented difficulties. For example argument (51) is intuitively valid.

(51) Every horse chased some cow.
    Every cow is brown.
    Every horse chased some brown cow.

One wants to say that the second premise licenses extension of ‘cow’ (in the first premise) into ‘brown cow’, even though ‘cow’ is part of a quantifier that is in turn part of a complex predicate that has been combined with a quantificational subject, along lines suggested by (51a).

(51a) [(every $\Phi$) is (one that $R$ some $\Psi$)]
    [(every $\Psi$) is $\Omega$]

where ‘$R$’ somehow corresponds to relations, but (51a) is not taken as a basic form of valid inference. But this highlights the need for a systematic way of accommodating relational notions, without spoiling the idea that quantifiers determine the inferential polarity of the predicates they combine with; see note (12) above.

By itself, (51) is not so troubling, since one can say that the second premise provides a general license to replace ‘cow’ with ‘brown cow. But consider (52).

(52) Every horse that saw some dog chased some cow that some yellow dog saw.
    Every horse that saw some yellow dog chased some cow that some dog saw.

Each occurrence of ‘dog’ appears in a constituent of the form ‘some $\Phi$’. Though in the first case, ‘dog’ can be validly replaced with the extended predicate ‘yellow dog’; here, the default direction of inferential polarity has been inverted. And of course, (52) is asymmetrically valid, in that its premise does not follow from its conclusion. But it’s not that the “main” quantifier determines inferential polarity, as shown by the equally valid (53) and (54).

(53) Some horse that saw every dog chased a cow that every dog saw.
    Some horse that saw every yellow dog chased a cow that every yellow dog saw.

(54) Every horse that saw no yellow dog chased a cow that no dog saw.
    Every horse that saw no dog chased a cow that no yellow dog saw.

Note that in the premise of (54), ‘yellow dog’ has the default inferential polarity—predicate reduction is licensed—as if the inverting effects of ‘no’ and ‘every’ cancel out.

To account for such facts, as opposed to merely listing them, one needs to do more than simply allow for predicates of the form ‘$\Phi$ that $R$ no $\Omega \Psi$’ and ‘$R$ some $\Delta$ that no $\Phi R$’. Even bracketing the further complexities illustrated with (55),

(55) Every farmer who owns a horse feeds it, himself, and his children.

---

16 It is worth remembering that Frege disagreed with Kant about the nature of logic. They agreed that arithmetic could not be reduced to Aristotelian syllogisms.
we need a systematic account of how predicates can have quantifiers and relational concepts as constituents, in ways suggested by the relative clauses in (52-54). One way or another, this requires expansion of the classical conception of thoughts. Predicates can be complex, in ways that are not merely conjunctive, without containing complete thoughts as constituents. Relational concepts can figure systematically in ways of generating cognitively productively predicates.

3. Composition by Adicity
Frege (1879, 188x, 1892a) helped a lot. But it is worth going slowly here—even, and perhaps especially, for those who have seen Frege’s notation used to theorize about words—because Frege was offering a proposal about the concepts and overt symbolism we ought to use in scientific inquiry. This proposal was intimately connected with Frege’s attempt to reduce arithmetic to a logic that treated monadicity as a special case of polyadicity involving relations that countable things can bear to truth values; see chapter two. So as Frege stressed, ordinary concepts/words need not and apparently do not conform to his ideal typology. One can speculate that pace Frege, his proposal was prescient cognitive science. My own view is that while Frege’s insights about conceptual architecture deserve a central place in psychology, we can live without truth values as relata, and natural lexical items do not exhibit Fregean types. Though in any case, theorists need to distinguish (i) claims about the concepts that are naturally available to humans, without special training or the invention of formal languages like Frege’s Begriffsschrift, from (ii) claims about the concepts that theorists ought to deploy for certain intellectual purposes.

3.1 Thoughts, Saturaters, and Unsaturates
One of Frege’s leading ideas was that complex concepts, including thoughts as special cases, can be analyzed as the results of saturating some concepts with others. Saturated concepts vary, as do their saturaters. But denoting concepts and complete thoughts cannot be saturated, while general and quantificational concepts are saturatable. To encode this crucial asymmetry, between thought components that get saturated and those that saturate others, Frege used notation that was designed to represent functions and their arguments: a monadic concept like \( \text{cow}(\_ \_ \_ \_) \) can be used to think about some things in a general way, while a denoting concept like \( \text{bessie} \) can be used to think about something as a particular; and combining \( \text{cow}(\_ \_ \_ \_) \) with \( \text{bessie} \) yields a complete thought, \( \text{cow}(\text{bessie}) \), according to which the individual denoted is one of the things thought about more generally. One can also view \( \text{cow}(\_ \_ \_ \_) \) as a “mental sentence frame” that is the result of abstracting away from the singular concept in \( \text{cow}(\text{bessie}) \).

Let’s say that thoughts are representations of a special type \( <t> \), and that denoting concepts are “entity-labels” of type \( <e> \). Monadic concepts can then be described as instances of type \( <e, t> \)—unsaturated representations that can be “filled” by a singular concept to form a complete thought. Since \( \text{cow}(\text{bessie}) \) is of type \( <t> \), \( \text{cow}(\_ \_ \_ \_) \) is of type \( <e, t> \). As stressed below, this does not yet require appeal to truth values, much less the Fregean idea that each ideal thought denotes a truth value. The typology here concerns concepts, not any posited denotations of ideal concepts. I assume that the concept \( \text{vulcan} \) can be of type \( <e> \), without there being any entity such that the concept applies to it, and that \( \text{cow}(\text{bessie}) \) can be of type \( <t> \) without there being any entity such that the thought applies to it.

---

17 The terminology in this field (intentional/intensional, concept/conception, etc.) conspires against us. Frege used ‘Begriff’ (‘Concept’) to talk about functions from entities to truth values; see §1 of chapter two. And given his concern with ideal representables, rather than actual representations, his notion of analysis was partly normative; cp. Quine’s (1963) notion of regimentation. But in my view, and I believe Frege’s, natural concepts can exhibit a saturated/unsaturated distinction without the unsaturated concepts actually representing (or being) Begriffs.
We can go on to describe a concept of negation that can be saturated by one thought to form another: \( \neg [\_], \) of type \( <t, t> \). Such a concept is formally monadic; though intuitively, it serves to classify thoughts with regard to correctness, rather than classifying entities in terms of some property or condition like being a cow. For example, \( \neg [\text{cow}(\text{Bessie})] \) is to be classified as true if and only if \( \text{cow}(\text{Bessie}) \) is to be classified as false. We can also describe a concept of existential closure, \( \exists [\Phi(\_)] \), that can be saturated by a monadic concept to form a complete thought. Here, \( \Phi(\_\_) \) is a placeholder for concepts of type \( <e, t> \). So \( \exists [\Phi(\_)] \) is a concept of type \( <<e, t, t>, t> \) that serves to classify monadic concepts according to whether or not they apply to something. For example, \( \exists [\text{cow}(\_)] \) is to be classified as true if and only if \( \text{cow}(\_\_\_) \) is such that it applies to something—i.e., if there is at least one cow. So \( \exists \) does not correspond to the classical \textit{SOME}QUANT; it is more like \( \text{SOME}QUANT \text{THING}_{\text{PRED}}_{\text{SUBJ}}^{18} \)

I’ll return to the role of variables in representations like \( \text{cow}(x) \) and \( \exists x[\text{cow}(x)] \). But let’s distinguish the mere idea of an unsaturated conceptual slot, as it figures in a Fregean taxonomy of concepts, from the idea of a slot that is linked to another via some variable.

### 3.2 Encoding Polyadicity

I am rehearsing familiar claims about concepts with meaning aforesight. It is widely held that human linguistic meanings recapitulate key aspects of the Fregean typology. For example, names are often treated as expressions of type \( <e> \), as if ‘Bessie’ is a pronounceable denoting concept; and the semantic effect of combining a name with a predicate is often described in terms of saturation. By contrast, I will eschew the usual semantic typology for linguistic expressions, and describe meaning composition in terms of conjunction. But I grant that thinkers often saturate one concept with another. Indeed, on my view, humans use concepts of various Fregean types—including polyadic concepts—to introduce conjoinable monadic concepts. And since I happily appeal to polyadic concepts, I need to represent them.

Frege offers two ways of representing a polyadic concept: as a concept that can be \textit{fully} saturated by one ordered \( n \)-tuple of singular concepts; or as a concept that can be \textit{partly} saturated by one denoting concept. For many purposes, this distinction does not matter. But it is worth being explicit about both options, especially in the context of explaining and evaluating the hypothesis that human linguistic meanings are concept assembly instructions.

Let’s say that the dyadic concept \textit{above}(\_, \_) can be used to think about one thing being above another, and that this concept can be partly saturated by a singular concept to form a complex concept like \textit{above}(\_, \text{Bessie}); where this monadic concept can be used to think about things above Bessie. The dyadic concept is of type \( <e, <e, t>> \); saturating it with a concept of type \( <e> \) yields a concept of type \( <e, t> \). Let \textit{above}(\_, \_) be a concept that can be saturated by an \textit{ordered pair} of denoting concepts to form a thought like \textit{above}((\text{Venus}, \text{Bessie})). Here, we needn’t worry about whether such saturation requires a \textit{concept of ordered pairs}. The idea is that \textit{above}((\text{Venus}, \text{Bessie})) has a monadic constituent whose saturater is complex in a way that

---

18. While denoting concepts cannot be saturated, predicative concepts can saturate “higher order” concepts of type \( <e, t, t>, t> \). I use regular angle brackets for conceptual types and steeper brackets for ordered pairs. So \( <e, t> \) is an ordered pair of types that can be identified with the set \( \{<e>, <e>, t>, t>\} \), just as \( \{1, 1, 0\} \) can be identified with the ordered pair of numbers \( (1, 0) \). One can describe a function \( F \) with a domain/range of two entities, \( T \) and \( \perp \) : \( F(T) = \perp \) and \( F(T) = \perp \). One can also stipulate that \textit{true} and \textit{false}, special concepts of type \( <e> \), denote \( T \) and \( \perp \). Then \( \text{tvC}(\_\_) \) can be a “truth value conversion” concept of type \( <e, e> \) such that: \( \text{tvC}([\text{true}]) \) is a way of thinking about \( \perp \), and \( \text{tvC}([\text{false}]) \) is a way of thinking about \( T \). But if natural thoughts are not concepts of truth values, then \( \neg [\text{false}] \) is conceptual gibberish, as is \( \text{tvC}([\text{cow}(\text{Bessie})]) \). For while \( \neg [\_\_] \) and \( \text{tvC}(\_\_) \) are both negation-concepts, the first requires saturation by a thought, and the second requires saturation by a concept of a truth value.
reflects the relational character of the thought. We can say that \( \text{above}(\_ , \_ ) \) is of type \(<e^2, t>\). But whatever our notation, we can stipulate that the formally distinct concepts are cognitively equivalent: for any concepts \( \alpha \) and \( \beta \) of type \(<e>\), \( \text{above}(\alpha, \beta) \iff \text{above}(\alpha, \beta) \).

Still, for some purposes, the formal distinction might matter. One can imagine a mind that can assemble the thought \( \text{above}(\text{venus}, \text{bessie}) \) without being able to form \( \text{above}(\_ , \_ , \_ ) \). Such a mind might need to saturate \( \text{above}(\_ , \_ ) \) all at once, with an ordered pair of concepts, because it cannot generate partly saturated structures like \( (\_ , \_ ) \). One can also imagine a thinker who can form \( \text{above}(\_ , \_ , \_ ) \) and \( \text{above}(\text{venus}, \_ ) \): concepts that can be used to think about, respectively, things above Bessie and things Venus is above. This is the kind of thinker Frege had in mind. Likewise, Frege was not worrying about a mind that could form \( \text{above}(\_ , \_ , \_ ) \) but not \( \text{above}(\text{venus}, \_ ) \). For his idealized purposes, \( \text{above}(\_ , \_ ) \) and \( \text{above}(\_ , \_ , \_ ) \) count as the same concept depicted in two slightly different ways: one that highlights the relational character of the concept; and one that highlights the possibility of representing relationality in terms of saturation by \( n \)-tuples. But if part of our task is to explain the unbounded productivity of human concepts, then the human capacity to form concepts like \( \text{thing-venus-is-above} \) is one of our explananda. So we need to think about the operations required to build concepts like \( \text{above}(\text{venus}, \_ ) \) from atomic constituents.

Perhaps we can build \( \text{above}(\text{venus}, \text{bessie}) \) and extract Bessie. But then one wants to know why the result is a saturatable representation that can be used to think about things that Venus is above, as opposed to mere residue of a destructive operation. (Extracting metals from ores, for example, leaves a mess.) I grant that concept abstraction is often possible. But it seems unlikely that every potential extraction/abstraction from \( \text{above}(\text{venus}, \text{bessie}) \) yields a concept that humans can naturally form: \((\text{venus, bessie}) \) would be of type \(<e, <e, t>, t>\); \((\text{venus, _}) \) would be of type \(<e, <e, t>>, <e, t>> \) or \(<e, <<e, <e, t>>, t>>\). One can stipulate that any thinker who has \( \text{above}(\_ , \_ ) \) and \( \text{venus} \) can form \( \text{above}(\text{venus}, \_ ) \). But then it is an empirical question whether ordinary humans acquire concepts like \( \text{above}(\_ , \_ ) \), as opposed to simulacra that allow for a more limited kind of composition/abstraction.

This point multiplies given other polyadic concepts. The technical and empirical details, which turn out to be interesting, are for later chapters. But I assume that humans and other animals have triadic concepts like \( \text{give}(\_ , \_ , \_ ) \) and \( \text{between}(\_ , \_ , \_ ) \) or \( \text{give}(\_ , \_ , \_ ) \) and \( \text{between}(\_ , \_ , \_ ) \)—that can be used to think about things that exhibit a certain ternary relation. Such concepts can be described as instances of type \(<e, <e, t>>, <e, t>> \) or \(<e, <<e, <e, t>>, t>>\). Somehow, humans can form concepts like \( \text{thing-venus-gave-to-bessie, thing-venus-gave-bessie-to} \), and \( \text{thing-that-gave-bessie-to-venus} \). We can even think about pairs of things such that the first gave Bessie to the second. Ignoring tense for simplicity, we can form concepts equivalent to: \( \text{give(venus, _, bessie)} \); \( \text{give(venus, bessie, _)} \); \( \text{give(_, bessie, venus)} \); \( \text{give(_, bessie, _)} \); etc. But children may start with composable concepts that are less cognitively productive.

Perhaps even adults have only limited capacities for partially saturating natural concepts. We may have \( \text{venus, bessie, and give(_, _, _)} \) but no capacity form to \( \text{give(venus, _, bessie)} \). We may instead have a capacity to generate a formally distinct concept—built conjunctively from concepts like \( \text{give(_, _)} \) and \( \text{agent(_, venus)} \)—that is equivalent for many purposes. And we may think about one thing giving a second to a third without thinking about the three things as exhibiting a ternary relation. But this is getting ahead. In this chapter, my goal is to explain and motivate the idea that phrasal meanings are instructions for how to build mental predicates, which can also be described as monadic concepts. So let me return to discussing adicity.
3.3 Analysis, Polyadicity and Junction
Among natural concepts, there may be a highest usable adicity. I have no idea what the limit is. But tetradic concepts may be common. Think about the vertices of a square, or about the difference between giving and selling. One can speculate that the concept \( \text{SELL}(\_, \_, \_, \_) \)—with a slot for what a buyer transfers to a seller in exchange for a thing bought—is complex, and that its component parts all have lower adicities. One can also claim that \( \text{TRANSFER}(\_, \_, \_) \) is analyzable in terms of dyadic and monadic concepts of causation, things, events of motion, and relational states of “having” things. But I don’t know how to provide such analyses.\(^{19}\)

On the other hand, we face unpleasant questions if the verb in ‘Venus sold Bessie’ expresses a tetradic concept. Most saliently, how can ‘sold’ combine with two grammatical arguments to form a sentence that can be used to express a complete thought? Part of my goal is to answer such questions, without positing covert grammatical arguments, by saying that humans use concepts like \( \text{SELL}(\_, \_, \_, \_) \) to introduce monadic concepts of events like \( \text{SELL}(\_) \), which can be combined with concepts like \( \text{AGENT}(\_, \_) \) and \( \text{PATIENT}(\_, \_) \); where these dyadic concepts, used to think about participation relations that events can bear to their participants, can be linked to grammatical relations that verbs can bear to subjects and objects. This raises questions, addressed below, about how concepts like \( \text{SELL}(\_) \) and \( \text{PATIENT}(\_, \_) \) are combined. But in my view, we should resist the idea that \( \text{SELL}(\_, \_, \_, \_) \) is composed of concepts that have lower adicities, and say that \( \text{SELL}(\_) \) is introduced in terms of concepts that have higher adicities. Likewise, I claim, for \( \text{GIVE}(\_, \_, \_) / \text{GIVE}(\_) \) and \( \text{CHASE}(\_, \_) / \text{CHASE}(\_) \).

As Frege recognized, introducing atomic symbols can be useful when symbols can be manipulated in ways that depend on their formal properties. Correlatively, “fruitful analyses” can reflect introduction of formally useful concepts, as opposed to decomposition of extant concepts; see Hory (2007). In chapter two, I discuss Frege’s illustration in terms of arithmetic concepts like \( \text{NUMBER}(\_) \), \( \text{THREE} \), and \( \text{PRECEDES}(\_, \_) \). But utility depends on what you want. Frege wanted to reduce arithmetic to logic. So he invented a new logic that was geared for relation notions. For other purposes, one might want to exploit a natural logic of predicates.

Suppose that \( \text{BROWN}(\_) \) can be joined with \( \text{COW}(\_) \) to form \( \text{BROWN}^\bowtie \text{COW}(\_) \), a complex predicate that can combine with a subject to form a thought while being used to think about things as cows and as brown. Imagine the unfilled slot of \( \text{BROWN}(\_) \) being identified with that of \( \text{COW}(\_) \), making the resulting junction monadic.\(^{20}\) Call this mode of combination “M-junction.”

---

19 Like Fodor and Lepore (2002), I think extant proposals fail if offered as decompositions of natural concepts. Even if transferring \( y \) to \( z \) is a matter of causing \( y \)’s possession of \( z \), I don’t think humans have a concept \( \text{DIFF} \) such that the concept with which we think about transferring \( y \) to \( z \) has constituents that include \( \text{CAUSE}, \text{POSSESS}, \) and \( \text{DIFF} \). One can posit a concept that distinguishes transfers from mere causings-to-have. But this is not yet to provide a plausible analysis, much less a plausible adicity-reducing analysis. And as we’ll see in chapters two and five, accounting for “one-way” implications (e.g, kill/die) does not require decomposition if concepts can be introduced.

20 If it helps, think of \( \text{BROWN}^\bowtie \text{COW}(\_) \) as the result of adjoining \( \text{BROWN}(\_) \) to \( \text{COW}(\_) \). In any case, \( ^{\bowtie} \) does not signify mere concatenation of mental symbols. To be a complex concept, \( \text{BROWN}^\bowtie \text{COW}(\_) \) has to differ from \( \{ \text{BROWN}(\_), \text{COW}(\_) \} \) and be relevantly like \( \text{COW}(\_) \): saturatable by Bessie, a potential saturater of \( \exists \Phi(\_) \), and joinable with other concepts. Kant and Frege stressed that a thought is not a mere list of ideas, echoing Plato’s observation that facts are not mere lists of objects. Likewise, a junction is not a mere list of concepts. Indeed, the notation leaves room for the idea that \( \text{BROWN}^\bowtie \text{COW}(\_) \) is the result of saturating a concept of conjunction, as in \(+\text{BROWN}(\_), \text{COW}(\_)\); where \(+\Phi(\_), \Psi(\_)\) is a concept of type \( <<e, t>, <<e, t>>, <e, t>> \). This would make conjunction an “indirect” mode of conceptual combination that operates via the “direct” mode of saturation. But one can also imagine a primitive joining operation that yields \( \text{BROWN}^\bowtie \text{COW}(\_) \), given \( \text{COW}(\_) \) and \( \text{BROWN}(\_) \) as inputs; where \( \text{COW}(\_) \) can be derived from \( \text{BROWN}^\bowtie \text{COW}(\_) \) in “default” contexts, as discussed in section two. Similarly,
without yet assuming that humans employ it. Given some atomic monadic concepts, \(M\)-junction can generate endlessly many more. But this kind of recursion is pretty boring. As noted above, the number of atomic monadic concepts puts a bound on the cognitive productivity of conjoining them. Generating \(COW^\wedge BROWN^\wedge COW(\_\_\_\_)\), and so on, is a cheap trick that earns little applause.

Let me note, digressively, that mere recursive saturation is similarly soporific:

\[\sim \{COW(BESSIE)\}, \sim \{\sim \{COW(BESSIE)\}\}, \sim \{\sim \{\sim \{COW(BESSIE)\}\}\}, \text{etc.}\] Concepts like \(\text{DENY}[\_\, \_\_]\) saturatable by a thought like \(COW(BESSIE)\) and a singular concept like \(VENUS\) to form a complex thought like \(\text{DENY}[VENUS, COW(BESSIE)]\), according to which Venus denied that Bessie is a cow—do yield a kind of unbounded cognitive productivity: \(\text{DENY}[BESSIE, COW(VENUS)]\), \(\text{DENY}[VENUS, \text{DENY}[BESSIE, COW(VENUS)]]\), etc. But human thought is unboundedly cognitively productive even setting aside concepts with which we think about relations that thinkers bear to representations. And a capacity to saturate concepts does not yet explain how we form concepts like \(COW\text{-THAT-CHASED-VENUS}, COW\text{-VENUS-CHASED}, COW\text{-THAT-DENIED-VENUS-IS-A-COW}, COW\text{-THAT-VENUS-GAVE-TO-BESSIE}, \text{etc.}\) Intuitively, construction of these concepts requires conjunction of \(COW(\_\_\_)\) with complex monadic concepts: \(\text{CHASE}(\_\_\_, \text{VENUS})\); \(\text{CHASE}(\_\_\_, \_\_\_)\); \(\text{DENY}[\_\_, \text{COW}(\_\_\_, \_\_\_)\]); \(\text{GIVE}(\_\_, \_\_, \text{BESSIE})\); etc. But to reiterate a point from §2.2, part of the task is to explain the human capacity to form concepts like \(THING\text{-VENUS-GAVE-TO-BESSIE}\).

One can say that given any polyadic concept with \(n\) unsaturated “slots,” a human thinker can use \(n\) saturated constructors to create a monadic concept, leaving any one of the slots unfilled. But that leaves the question of how we came to have this impressive capacity. And in chapter five, I offer evidence that a simple form of conjunction lies at the core of unbounded cognitive productivity. Our natural capacities to combine concepts are impressive, but constrained in ways that suggest less than an ideal Fregean mind. I think we make clever use of monadic concept conjunction, as medieval logicians suspected, along with a few other tricks: the introduction of concepts like \(\text{GIVE}(\_\_)\); appeal to thematic concepts like \(\text{AGENT}(\_\_, \_\_)\), \(\text{PATIENT}(\_\_, \_\_)\), and \(\text{RECIPIENT}(\_\_, \_\_)\); a very limited operation of existential closure; and one other operation, described in chapter six, that is required on any account. For specialists, let me say at once: the additional operation is a mental analog of relative clause formation, but much less powerful than lambda abstraction, and limited to monadic concepts of a special sort. The idea will be that with help from human language syntax, we can generate an analog of \(\text{GIVE}(\text{VENUS}, \_\_, \text{BESSIE})\) without saturating \(\text{GIVE}(\_\_, \_\_, \_\_)\)—much less saturating it twice, or thrice and then desaturating once.

Before the digression/advertisement, we were considering the old idea that thinkers can join concepts, thereby forming conjunctions like \(BROWN^\wedge COW(\_\_\_)\). And to recapitulate: a mere capacity to join monadic concepts cannot deliver \(THING\text{-VENUS-GAVE-TO-BESSIE}\); though neither can a mere capacity to saturate concepts, without a further capacity (that humans may not have) to extract any argument from a multiply saturated concept. Nonetheless, I think a variant of the old idea can help explain both the cognitive productivity of human concepts and the semantic character of human linguistic expressions, at least if the variant is embedded in a broader proposal according to which meanings are concept assembly instructions.

4. Add a Little Dyadic
To get beyond \(M\)-junction, which can only generate concepts like \(BROWN^\wedge COW(\_\_\_)\), I posit a joining operation that allows for conjunction of concepts like \(\text{PATIENT}(\_\_, \_\_)\) and \(COW(\_\_)\). But the posited operation allows for only a smidgen of polyadicity. An atomic dyadic concept can be saturating one concept with another does not require a concept of saturation. But for now, it doesn’t matter if \(BROWN^\wedge COW(\_\_\_)\) really has three conceptual constituents, like \(+[BROWN(\_\_), COW(\_\_\_)].\)
combined with a monadic concept subject to a severe constraint: the first slot of the dyadic concept must link to the monadic concept, which must be existentially closed as shown below.

\[ \exists [\text{PATIENT}(_, _) \& \text{COW}(\_)] \]

This yields a complex monadic concept whose sole open slot is the second slot of the dyadic concept—counting from the right, so that the first (or “most internal”) slot of any predicative concept is the one that would be saturated first. Recall that saturating \text{PATIENT}(_, _) with \text{BESSIE} would yield \text{PATIENT}(\_, \text{BESSIE}), which is a concept of things that have Bessie as their patient. Likewise, the concept indicated above applies to something if and only if its patient is a cow.

### 4.1 Very Limited Conjunction

The idea is to characterize a form of conjunction that recursively generates monadic concepts, and no others, but temporarily tolerates an atomic dyadic conjunct. In terms of the example: the first slot of \text{PATIENT}(_, _) is linked to the adjacent monadic conjunct, as if two monadic concepts were M-conjoined; but the presence of a second slot requires closure of the first, since conjunctions must be monadic. Put another way, while the posited system permits conjunction of an atomic dyadic concept with a complex (i.e., generated) concept, this system cannot conjoin two polyadic concepts of any kind. So no generated conjunct can be dyadic.

As discussed in section three, conjoining \text{CHASED}(\_, \_) with \text{SAW}(\_, \_) would require a determination of how the four conceptual slots are to be linked or left unlinked. And it is all too easy to stipulate that polyadic concepts can be combined in productive ways, as if humans are born with ideal Fregean minds. So I want to start by characterizing a recursive form of conjunction that generates monadic concepts, in a way that goes beyond M-junction, but without assuming that polyadic concepts can be conjoined subject to further constraints that need to be theoretically motivated and biologically implemented. More specifically, I want to ask if a variant of M-junction that allows for a smidgeon of polyadicity can serve as the core recursive operation of semantic composition for human languages and thereby help explain human cognitive productivity. To operationalize ‘smidgeon’, I am restricting polyadicity to a single dyadic conjunct that must be (i) atomic rather than generated, and (ii) immediately made part of a complex monadic concept, as opposed to leaving both slots open to further processing.

Call the envisioned form of conceptual combination “\text{\textTheta}-junction,” as a reminder that the dyadic concept can be thematic and indicated by a grammatical relation.\textsuperscript{21} For some purposes, it would do no harm to represent results of \text{\textTheta}-junction with an ampersand and variables, as in \( \exists \text{[PATIENT}(\text{E}, \text{X}) \& \text{COW}(\text{X})] \). But as we’ll see, given standard interpretations for the ampersand and variable, the concept \( \exists \text{[PATIENT}(\text{E}, \text{X}) \& \text{COW}(\text{X})] \) has sophisticated constituents that support a kind of cognitive productivity that ordinary humans may not enjoy. Note that \( \exists \text{[COW}(\text{X}) \& \text{CHASED}(\text{X}, \text{VENUS}) \& \text{GIVE}(\text{VENUS, X, BESSIE}) \& \text{BROWN}(\text{BE} \text{SSIE})] \) may not be a natural concept. So the linking notation, used above, is not a mere typographic alternative to using variable letters and an ampersand. The hypothesis is that humans can employ a certain mode of concept junction that permits a little more than M-junction. The links indicate how the posited form of junction

\textsuperscript{21} One can characterize other “minimally nonmonadic” forms of junction. But \text{\textTheta}-junction has an obvious appeal for neo-Davidsonian semanticists who regularly appeal to event predicates and thematic relations. Cp. Higginbotham’s (1985) discussion of theta-binding and theta-linking.
(deterministically) accommodates a dyadic input. The links do not indicate a particular choice about how to use three independent conceptual slots in forming a complex concept.

I’ll say more about this, and the posited form of closure, in section five. But since the linking notation is typographically inconvenient, let me depict the Θ-junction of COW(_ _) with PATIENT(_, _) as follows: \(\exists[\text{PATIENT(_, _) \text{COW}(_ _)}]\). Despite the absence of variable letters, this is unambiguously a monadic concept that applies to an entity—with events as special cases of entities—if and only if that entity has a cow as its patient: the first/rightmost slot of PATIENT(_, _) must link to the slot of COW(_ _), which must be the target of closure. In general, the Θ-junction of a monadic concept \(\Phi(_)\) with a dyadic concept \(\Delta(_, _)\) is \(\exists[\Delta(_, _) \Phi(_ _)]\), which applies to an entity if and only if \(\Delta(_, _)\) applies to the pair consisting of that entity and something to which \(\Phi(_)\) applies.

One can describe Θ-junction as an operation that yields a concept of type \(<e, t>\) given inputs of type \(<e, <e, t>>\) and \(<e, t>\). But this way of encoding semantic typology is misleading if all nonmonadic inputs to composition operations are dyadic and atomic. Even if one takes Θ-junction to be a special case of saturation, since a dyadic concept combines with something to yield a monadic concept, this operation does not support “adicity reduction” for concepts of any other type. Again, Θ-junction does not permit supradyadic conjuncts like BETWEEN(_, _ _) or conjunctions of dyadic concepts. Indeed, Θ-junction does not even permit junction of PATIENT(_, _) \text{COW}(_ _), with a monadic concept: VIOLENT(_ _)[\text{PATIENT(_, _) \text{COW}(_ _)}] is not generable. The envisioned system cannot create a complex dyadic concept by linking the unsaturated slot of VIOLENT(_ _) to one of the other slots; nor can it create a triadic concept by leaving the new slot unlabeled. Only a little dyadicity is allowed, and none that percolates.

One can view ‘^\text{^}\text{^}' and ‘\exists' as together indicating a conjunctive operation that can be described as bipartite, leaving it open whether the implemented operation is divisible into junction and closure components. Humans may enjoy a slightly flexible mode of junction that always yields a monadic concept, but one whose form depends on the inputs: given \(\Phi(_)\) and \(\Psi(_ _),\) the result is \(\Phi^\text{^}Psi(_ _),\) in which the slot of \(\Phi(_)\) is identified with the slot of \(\Psi(_ _);\) given \(\Delta(_, _)\) and \(\Phi(_ _),\) the result is \(\exists[\Delta(_, _) \Phi(_ _)].\) Alternatively, we may enjoy two related forms of junction, each type-inflexible. (In which case, using ‘^\text{^}\text{^}' for both M-junction and Θ-junction could be misleading.) It is hard to tell the difference, in part because one might analyze concepts like BROWN-COW as Θ-junctions like \(\exists[\text{BROWN}(_, _) \text{COW}(_ _)],\) with BROWN(_ _) applying to an ordered pair \((\alpha, \beta)\) if and only if \(\alpha\) is: brown and identical to \(\beta;\) or the predominantly brown surface of \(\beta;\) or identical to \(\beta\) and brown for a thing of the specified sort; etc.

With those caveats entered, I’ll use ‘^\text{^}\text{^}' to encode both M-junction and Θ-junction. Thus, \(\exists[\text{PATIENT(_, _) \text{BROWN} \text{COW}(_ _)}]\) applies to an entity if and only if that entity has a brown cow as its patient. Likewise, \(\text{CHASE}(_) \exists[\text{PATIENT(_, _) \text{COW}(_ _)}]\) applies to an entity if and only if it is a chase that has a cow as its patient. And to flag a more interesting case, discussed later,

\(\exists[\text{PATIENT(_, _) \text{ARRIVAL}(_) \exists[\text{PATIENT(_, _) \text{BROWN} \text{COW}(_ _)}]]]\)

applies to an entity if and only if it has an arrival whose patient is a brown cow as its patient.\(^22\)

If one wants to make the contrast between M-junction and Θ-junction more vivid, one can

\(^{22}\) For simplicity of initial exposition, pretend that all “internal” arguments of verbs correspond to \(\text{PATIENT(_, _)}\). Chapters five and six address the complications concerning which thematic concepts correspond to which verbs. For example, things that arrived may be themes of arrivals, even if things kicked/chased are patients of kicks/chases. And arrivals that are seen may be themes of perceptual events that have experiencers rather than agents.
recode M-junctions as follows: Brown^cow(\_). But I'll want to highlight a potential parallel with predicate-adjunct and predicate-argument combination in human language syntax: while adjunction differs from complementation, expressions are combined in both cases, suggesting some common mode of syntactic combination that can be supplemented; see Chametsky (19xx), Hornstein and Pietroski (2009), Hunter (2011). Let me defer such details, though, and finally get back to clarifying the idea of concept assembly instructions.

4.2 Fetch, Join, Repeat

If biology somehow implements M-junction and Θ-junction, one can envision a mind with further capacities to (i) use lexical items to access concepts that can be inputs to these operations, and (ii) combine lexical items in ways that trigger assembly of concepts. Suppose that a “merging” two expressions (lexical or phrasal) is an instruction to send a pair of corresponding concepts to a “joiner,” whose output can be an input to further operations of joining, perhaps via storage in some form of short term memory. And suppose that a second way of combining two expressions, “thematizing,” sends a corresponding pair of concepts—one monadic but perhaps complex, the other dyadic and atomic—to the Θ-joiner, whose monadic output can also be an input to further operations. This would yield a kind of recursion that goes beyond M-junction.

As a concrete illustration, imagine the following sequence of operations.

1. fetch@`cow`
2. fetch@`brown`
3. M-join[1, 2]
4. fetch@Θ:pat
5. Θ-join[3, 4]
6. fetch@`chase`
7. M-join[5, 6]

Let the numbers correspond to reusable memory registers, where results of executing instructions can be stored, and let Θ:pat be a functional item whose address is that of patience(\_, \_). If see(\_) and arrival(\_) are stored at the lexical addresses for ‘see’ and ‘arrive’, then the monadic concept see(\_)∃[patience(\_, \_)][arrival(\_)∃[patience(\_, \_)][brown^cow(\_)]) can be similarly assembled by executing the following instruction.

M-join[fetch@`see`],
Θ-join[fetch@Θ:pat],
M-join[fetch@`arrive`],
Θ-join[fetch@Θ:pat, M-join[fetch@`brown`, fetch@`cow`]]

Executing this instruction yields a concept akin to but distinct from the more sophisticated
For these purposes, instructions can take many forms, perhaps even strings of ‘1’ s and ‘0’ s that get used—as in a von Neumann machine—to access other such strings and perform certain operations on them. The instruction ‘1101011010’ might be executed by performing operation six (110) on whatever number is stored in (1) register twenty-six (011010), while ‘0100110101’ calls for performing operation two (010) on the number (0) fifty-three (110101). And instead of arithmetic operations that are performed on accessible/generable numbers, one can imagine conjunctive operations that are performed on accessible/generable concepts.

I won’t appeal to instructions of the form Saturate[I, I’], which would call for saturating the result of executing one instruction with the result of executing another. But others can hypothesize that chase( , ) and Bessie are lexically fetchable, and that chase( , Bessie) is assembled by executing Saturate[fetch@‘chase’, fetch@‘Bessie’]. Likewise, every[Brown cow]—a concept of type <<e, t>, t>—might be assembled by executing Saturate[fetch@every’, M-join[fetch@‘brown’, fetch@‘cow’]]23 The idea of executing Begriffsplans, thereby assembling concepts, is neutral with regard to the atomic concepts and composition operations. But the availability of certain concepts/operations in cognition does not ensure the constructability or executability of corresponding instructions to fetch/combine. Imagine a mind that can saturate some concepts with others, conjoin certain concepts, etc.

If Brown( ) and Cow( ) are the only concepts stored at the corresponding lexical addresses, then even if the instruction Saturate[fetch@‘brown’, fetch@‘cow’] is generable, it will not be executable. Given the lexical limitation, ‘brown cow’ would not be an executable instruction for a mind that could only generate Begriffsplans of the form Saturate[I, I’]. Though of course, type shifting can extend the utility of a constrained instruction generator. Suppose that Brown( ) can be used to introduce Brown[ , Φ( ) ]; a concept of type <<e, t>, <e, t>> that when saturated by a monadic concept Φ( ) yields a concept that applies to an entity if and only if that entity falls under both Brown( ) and Φ( ). If Brown[ , Φ( )] is assigned to the lexical address for ‘brown’, along with Brown( ), then one way to execute fetch@‘brown’ is to fetch the higher-order concept Brown[ , Φ( )]. In which case, executing Saturate[fetch@‘brown’, fetch@‘cow’] could result in the assembly of Brown[ , Cow( )]—a concept equivalent to Brown<>cow( ).24

This kind of type shifting may or may not be natural for humans. So let’s also imagine a mind that can only generate assembly instructions of the form M-join[I, I’] or Θ-join[I, I’]. If chase( , ) and Bessie are the only concepts stored at the corresponding lexical addresses, then the instruction M-join[fetch@‘chase’, fetch@‘Bessie’] will not be executable, and likewise for Θ-join[fetch@‘chase’, fetch@‘Bessie’]. The concepts chase( , ) and Bessie cannot be joined. But suppose that chase( , ) and Bessie are used to introduce chase( ) and bessie( ); where to a first approximation, bessie( ) applies to an entity if and only if it is the entity thought about with bessie. If chase( ) and bessie( ) are assigned to the relevant lexical addresses, then executing M-join[fetch@‘chase’, Θ-join[fetch@‘Θ:pat’, fetch@‘Bessie’]] can be a way to assemble chase( )∃[Patient( ), bessie( )], a concept of chases of a/the Bessie. Similarly,

24 Cp. Parsons (1970), Kamp (1970), Montague (1974). This raises the question of whether disjunctive concepts could be introduced, so that ‘brown cow’ might be roughly synonymous with ‘is brown or is a cow’. Schein (2002, forthcoming) discusses related questions—illustrated with ‘John and Mary formed a happy union/make a good couple/are happy together’—in the context of an approach to plurality adopted in chapters three and four.
∃[AGENT( _, _ ) DOG( _ )]∃[CHASE( _ )]∃[PATIENT( _, _ ) COW( _ )] can be assembled by executing the instruction below, letting the address of ‘Θ:ag’ be that of the concept AGENT( _, _ ).

M-join[Θ-join[fetch@‘Θ:ag’, fetch@‘dog’],
     M-join[fetch@‘chase’, Θ-join[fetch@‘Θ:pat’, fetch@‘cow’]]]

There are many details to work out. But my suggestion is that human linguistic expressions pair instructions of this sort with pronunciations. Or put another way, human languages are (biologically implementable) procedures that generate expressions whose meanings are concept assembly instructions of a certain sort. Each generable expression has a phonological side (PHON) and a semantic side (SEM), by virtue of which the expression can interface with the articulatory/perceptual systems and intentional/conceptual systems that are employed in the production and comprehension of speech. The idea is that meanings/SEMs can be described as generable instructions for how to assemble concepts. Each meaningful and articulable expression can also be described syntactically—e.g., as a noun phrase headed by ‘cow’. Indeed, such description is crucial for theories of the generative procedure. So we want to know how syntactic descriptions are related to SEFs. But a noun phrase can be a concept assembly instruction, as well as an instruction for how to activate articulators. (Noun phrases can be analytically related; they can also rhyme.)

Meaningful expressions can be articulable Begriffsplans that have syntactic properties. But such expressions can also be syntactic structures that have meanings and articulations, since such structures can pair PHONS with SEFs. On this view, meanings can be as composite as the corresponding concepts. Meanings—are like phrases—are not merely specifiable in terms of lexical items and syntax. Instructions like M-join[fetch@‘brown’, fetch@‘cow’] have simpler instructions as parts. But individuals can share Begriffsplans without sharing concepts. Speakers might fetch different concepts with ‘cow’, or articulate this word in slightly different ways, yet still count (at least for many purposes) as pairing the same phonological instruction with the same semantic instruction. We may not have clear criteria for what counts as the same lexical address across speakers; and the vague facts may be determined in part by what fetched concepts are concepts of. But as suggested by polysemy, a single instruction may correspond to multiple concepts.

I hope the analogy to elementary computer programs, which can be compiled and implemented, makes the operative notion of instruction tolerably clear and unobjectionable in the present context. One can raise many questions about the sense(s) in which Begriffsplans are intentional, and which philosophical projects would be forfeited by appealing to unreduced capacities to generate and execute the instructions posited here. But my task is not to reduce linguistic meaning to some nonintentional basis. I’ll say more about Begriffsplans in later

---

25 This may require three memory registers, and likewise for Saturate[Saturate[fetch@‘every’, fetch@‘dog’], Saturate[fetch@‘chase’, fetch@‘Bessie’]]. However semantic composition is described, concept assembly can reflect complexity that is due to more than one slot of a polyadic concept. One expects limits on how much such complexity a natural system can tolerate within one cycle of computation; cp. Chomsky (2005), Boeckx (200x). But ‘Every dog that saw a cat chased a cow that saw a rat’ may well call for storage of at least one concept constructed by executing a “phase” of the expression; cp. Uriagereka (2002). So it seems safe to assume repeatable use of at least three memory sites in which concepts can be stored as inputs to a later stage of combination.

26 Following Chomsky (1964, 1965) and many others, I think the I-languages that generate complex expressions (from atomic expressions) can and should be described in purely formal/syntactic terms, without reference to the interfacing systems. So I suspect that human languages can be described without talk of instructions, phonological or semantic. But in describing a generative procedure as a language, one suggests—and I grant—that the procedure can be used in a way that supports communication.
chapters. But in offering an account of how human linguistic meanings combine, I assume concepts and capacities to generate and execute instances of certain instruction types: \textsc{M-join}[I, I’]; \textsc{Θ-join}[I, I’]; and a couple of others to be discussed. For even if meanings are Begriffssplans, semanticists face an amply hard two-part question: what types of concepts can lexical items call for, and what modes of concept composition do phrasal meanings call for? Working out and defending a relatively spare answer is task enough.

For example, since neither \textsc{M-junction} nor \textsc{Θ-junction} provides a way of forming complete thoughts, I need to say something about sentential conjunction, sentential negation, and the very idea of a matrix sentence. Relatedly, one might think that any plausible analysis of quantificational constructions (e.g., ‘Fido chased every cow’) will appeal to Fregean instructions to saturate concepts. I’ll try to show otherwise, drawing on Pietroski (2005a). More generally, I think that we can eliminate appeal to \textsc{Saturate}[I, I’] in favor of appeal to \textsc{Θ-join}[I, I’], and that we cannot avoid appeal to the latter. But in any case, accommodating relative clauses—as in ‘dog that chased a cow that chased a dog’, etc.—requires a further “abstractive” operation; see chapters two and six. This highlights a methodological point, with which I’ll end this section.

Relative clauses reflect a cognitively productive form of recursive combination. They also suggest that sentential expressions can be converted into phrases that correspond to monadic concepts that can be conjoined with others. In light of such clauses, one wants to know—even for sentences with no additional adjectives, adverbs, or prepositional phrases—(i) what forms of conjunction and abstraction will be required by any decent theory that also appeals to a form of saturation, and (ii) what forms of conjunction and abstraction are required by a decent theory that \textit{does not} appeal to a form of saturation. Given any posited operation of conceptual composition, one wants to know if it is needed to achieve the kind of productivity illustrated with ‘dog that chased a cow that chased a dog’. For if not, then one needs other reasons for supposing that the operation is invoked by a composition instruction. I propose a bundle of operations/instructions that does not include \textsc{Saturate}[I, I’]. On this view, lexical items can only call for monadic or dyadic concepts, because semantic composition is a matter of building monadic concepts from monadic or (atomic) dyadic inputs. We’ll see if this turns out to be plausible.

5. Flavors of Conjunction

At this point, let me enter a confessional note. For many years, I used ‘&’ to encode claims about logical forms without asking (much less knowing) what I intended to represent with that squiggle. I knew the ampersand of an invented predicate calculus, as in ‘Brown(x) & Cow(x)’, was not the one defined in terms of truth tables. And I suspected that understanding ‘brown cow’ did not require tacit grasp of Tarski’s (1933) notion of conjunction—characterized in terms of satisfaction of open sentences, which can have unboundedly many variables, by unboundedly long sequences of entities. But I had no clear alternative in mind when inscribing formulae like ‘\(\exists e[\text{Stab}(e, \text{Brutus}, x) \& \exists z[\text{With}(e, z) \& \text{Knife}(z) \& \text{Today}(e)]]\)’ or ‘\(\text{Stab}(e) \& \text{Patient}(e, x)\)’, in the course of talking about human languages. Mea culpa. My intention is to do better.

The posited concept \(\exists[\text{Patient}(\_ , \_)]^\forall[\text{Brown}(\_)^\forall\text{Cow}(\_)]\)—formed by \textsc{M-joining Brown}(\_ with \text{Cow}(\_), and \textsc{Θ-joining} the result with \text{Patient}(\_ , \_)—may be a figment of theoretical imagination. But it is not \(\exists[\text{Patient}(e, x) \& [\text{Brown}(x) \& \text{Cow}(x)]]\) in disguise. My proposed modes of junction may turn out to be unnatural, in that ordinary thinkers do not employ them. Though as we’ll see, Tarskian conjunction is unlikely to be more natural, despite its

---

27 I allow for further reduction without providing any. Perhaps \textsc{Θ-join} can be reduced to \textsc{M-join} and an instruction to perform a restricted kind of closure. And perhaps the basic instructions are intentional only in a limited sense.
distinguished history. And if only for these reasons, I want to locate M/Θ-junction within a range of possible conjunctive operations. Repeating a point from §2.1, encoding possible concept types helps in specifying which types correspond to human linguistic meanings. A typology of conjunctions will also help clarify my own proposal, indicate some other options, and further illustrate the point that Θ-junction allows for only a smidgeon of polyadicity.

5.1 Propositions and Predicates
For simplicity, let’s encode the options in terms of dyadic concepts of conjunction, and say that a conjunctive thought/concept is the product of saturating a concept of conjunction; see note 20 above. We can worry later, if we’re not too tired, about converting talk of conjunction-concepts into other terms. We have enough to think about now, since in principle, there are endlessly many conjunction-concepts. But let’s start with some obvious examples that allow for recursion.

Let &(_ , _) be a concept of type <t, <t, t>> that can be saturated twice, by a thought each time, to yield a complex thought that is true iff both saturating thoughts are true. For these purposes, ignore any differences between &(_ , _) and two formally distinct concepts: &(_ , _), the corresponding concept of type <t', t> that can be saturated by an ordered pair of thoughts; and &(_ , _), the corresponding concept that can be saturated by an unordered pair of thoughts. Then we can use ‘&’ to indicate conjunction of thoughts—representations of type <t>—abstracting from further formal details. Restricting the ampersand in this way leaves room for the hypothesis that thoughts like VENUS-ROSE-AND-BESSIE-CHASED-VENUS are formed via some other (perhaps more permissive) operation. But one can also hypothesize, not implausibly, that humans and other animals enjoy a conjunction-concept of type <t, <t, t>>.

Let ∩[Φ( ), Ψ( )] be of the type <<e, t>, <<e, t>, <e, t>>—or abbreviating ‘<e, t>’, the type <et, <et, et>>. When saturated by monadic concepts, ∩[Φ( ), Ψ( )] yields their M-junction. So the instruction below has the same effect as M-join[fetch@‘cow’, fetch@‘brown’].

Saturate[Saturate[fetch@‘∩’, fetch@‘cow’], fetch@‘brown’]

The notation reflects the old idea that such junctions represent intersections of properties. Again, ∩[COW( ), BROWN( )] differs formally from ∩[COW( ), BROWN( )] and ∩[COW( ), BROWN( )]. But let’s use the intersection sign to indicate any such conjunction of monadic concepts, without concern about order, leaving room for various hypotheses about if and when humans deploy concepts of type <et, <et, et>>; again, see note 20.

One might balk at this “higher” conceptual type. But &(_ , _) and ∩[Φ( ), Ψ( )] are both concepts of types that exemplify the pattern <F, <F, F>>; where F is a Fregean type. Given joins of monadic concepts, we have a thought-concept of type <t, <t, t>> does not seem like a big leap, and likewise for a junction-concept of type <et, <et, et>> given joinable monadic concepts. Appeal to higher-order relational concepts seems unavoidable, given thoughts like MOST-OF-THE-COWS-ARE-BROWN and FIDO-CHASED-MOST-OF-THE-COWS, often analyzed as follows: MOST[BROWN( ), COW( )], and MOST[CHASED(FIDO , ), COW( )]; where the quantificational concept MOST[Φ( ), Ψ( )] is of type <et, <et, t>>.  

28 In chapter six, I’ll say how a dyadic concept like MOST[Φ( ), Ψ( )] can be used to introduce a monadic analog that can be conjoined with others. For now, just note that unlike COW( ), MOST[Φ( ), Ψ( )] may well be analyzable. Indeed, there is more than candidate analysis. A concept of one-to-one correspondence—saturable with COW( ) and BROWN( ) to form ONE-TO-ONE[BROWN( ), COW( )], according to which the cows correspond one-to-one with the brown things—might be used to introduce ONE-TO-ONE-PLUS[Φ( ), Ψ( )], permitting construction of a thought according to which: some (but not all) of the cows correspond one-to-one with (all of) the brown things. One could then identify MOST[Φ( ), Ψ( )] with ONE-TO-ONE-PLUS[Φ( ), Ψ( )]. Or suppose we have a concept NUMBER[Φ( )] of type <et, e>; where here, italics just suggest concepts that let one think about numbers. Given GREATER(X, Y)—a
That said, $\cap[\Phi(\_), \Psi(\_)$ differs from $\&(\_, \_)$ in a way that merits attention: \textsc{cow}(\_) is a monadic concept with a “slot” of its own, and likewise for \textsc{brown}(\_). Using a link as below,

\[
\text{\textsc{brown}(\_) $\&$ \text{\textsc{cow}(\_)$]}
\]

or a variable letter as in ‘$\cap[\text{\textsc{cow}(x), \text{\textsc{brown}(x)}$]’, makes it explicit that the “two” slots count as occurrences of the same unsaturated position in the joined concept. Perhaps such linking is trivial for concepts of type $<e, t>$. But if only because biological \textsc{and}-gates for representations of type $<t>$ may be easier to build than monadic analogs, the distinction between propositional conjunction and (any variant of) monadic concept conjunction is worth noting.

More generally, having concepts of type $F$ does not ensure the availability of a conjunction-concept of type $<F, <F, F>>$. Perhaps $\text{\textsc{chase}(\_, \_)$ and $\text{\textsc{say}(\_, \_)$ can be joined to form a third concept of type $<e, et>$ that can be indicated with variable letters as follows:

\[
2 \cap[\text{\textsc{chase}(x, y), \text{\textsc{say}(x, y)}$]
\]

But there is no guarantee that such conjunction/linking is possible. We may not have a conjunction-concept of type $<<e, et>, <<e, et>$>; \textit{cp.} Montague (197x).

And even if we have a dyadic concept-conjoiner, which might be useful for various cognitive purposes, it may not be accessible for purposes of linguistic comprehension. Similar points apply, with more force, to supradyadic concepts like $\text{\textsc{give}(\_, \_ \_)}$ and $\text{\textsc{sell}(\_, \_ \_ \_)}$.

4.2 Synecologematic Conjunction

With this in mind, consider a third notion of conjunction, reflected in formal languages that allow for open sentences like ‘Rxy & Syzw’ and closed sentences like ‘Rab & Sbcd’. This italicized ampersand, unlike ‘&’, can connect sentences that contain arbitrarily many variables (treating constants as special cases); see chapter two. Thanks to Tarski (193x), this kind of conjunction has become part of the logician’s standard toolkit. But it has some striking features that may not be exhibited by any natural mode of concept junction.

Note that ‘Rxy & Syzw’ corresponds to a tetradic concept, even though each conjunct corresponds to a dyadic or triadic concept. Because this ampersand connects sentences that are spelled with variable letters, which determine how the predicative slots are linked, ‘Rxy & Syzw’ has a fixed adicity.29

But since there are no constraints on how variables can reoccur across sentences, the adicity of a Tarskian conjunction is not determined by the adicity of the conjuncts, except in special cases where at least one conjunct has zero variables. While each of the conjuncts in ‘Bx & Cy’ is monadic, the conjunction is dyadic as ‘Rxy’.

One can hypothesize that humans enjoy a concept of conjunction that takes a pair of “pre-linked thought-frames” as inputs; where a \textit{thought-frame} is any (complete thought or

\begin{footnotesize}
\footnote{By contrast, one could stipulate that ‘R\_\_ & S\_\_’ is well-formed but of no fixed adicity. Depending on how the slots get linked, this conjunction might be triadic, tetradic, or pentadic—or monadic or dyadic if there can be linking within a sentence that has no sentential constituents, as in ‘Rxx’.}
\end{footnotesize}
predicative concept, and thought-frames are pre-linked if their slots are marked—as with
variables—in a way that determines which are linked to which, and which are left unlinked.
Saturating such a concept (of conjunction) twice would yield another pre-linked thought-frame,
though one whose adicity might exceed that of either saturater. We can say that such a concept
would be of the type \(<\text{PLTF}, \text{PLTF}, \text{PLTF}>>\). But this is not a Fregean type: the first conjunct
might be of any type; likewise for the second; and the input types need not determine the output
type. One can say that the inputs are Fregean concepts like \(\text{ABOVE}(., .)\) and \(\text{GIVE}(., ., .)\) along
with a choice of how to link the slots. But this third input, the choice, would not be a saturater in
any intuitive sense. Correlatively, “saturating” a concept of type \(<\text{PLTF}, <\text{PLTF}, \text{PLTF}>>\) with
\(\text{ABOVE}(., .)\) and \(\text{GIVE}(., ., .)\) would not yield a monadic concept, saturatable by \(\text{BESSIE}\); it would
yield a representation with no fixed adicity (see note 29).

If we want to talk about a concept of conjunction that has these features, I think our
notation should highlight its distinctive character. It is potentially misleading to use the same
shape used for the relatively boring concept of propositional conjunction, which is of the fixed
type \(<\text{t}, \text{t}, \text{t}>>\). So let’s replace ‘&’ with ‘!’ and stipulate that \(\text{ABOVE}(X, Y)! \text{GIVE}(Y, Z, W)\) is a
concept with which one can, if one has the concept, think about one thing being above a second
that a third gave to a fourth. Given special training and a blackboard, one can invent and perhaps
think with such symbolism. But inventing a procedure that can generate ‘Ba!Cb’, ‘Bx!Cx’,
and ‘Rxy!Gyzw’ hardly shows that a natural procedure generates analogs of all these invented
conjunctions if it generates analogs of one.\(^{30}\) Assuming that human concepts are conjoinable in
this Tarskian way is to assume, rather than explain, a very strong form of cognitive productivity.

Questions about the available concept(s) of conjunction raise corresponding questions about
existential closure. The concept \(\exists[\Phi(.)]—\) of type \(<\text{et}, \text{et}>>\)—can be saturated by \(\text{ABOVE}(., \text{BESSIE})\)
or \(\text{ABOVE}(\text{VENUS}, .)\). But \(\exists[\Phi(.)]\) cannot combine with \(\text{ABOVE}(., .)\) or any other concept of type
\(<\text{e}, \text{et}>\); see chapter two for further discussion. We can introduce a more permissive concept of
existential closure that can combine with any predicative concept, and bind any variable, to allow
for triadic concepts like \(\exists z[\text{ABOVE}(X, Z)! \text{GIVE}(Y, Z, W)]\). Perhaps humans can use this concept to
think about a first thing being above something which a second thing gave to a third. But perhaps
in so far as we naturally do such thinking, we use concepts that exhibit a different kind of
architecture. Similar points apply to negation, since by hypothesis, \(\sim[.]\) is of type \(<\text{t}, \text{t}>>\).

Put another way, one can learn to see ‘\(\exists y(\text{Axy} \& \text{Gyzw})\)’ as an expression that is satisfied by
certain sequences of things—in accord with Tarski’s algorithm, discussed in chapter two. And
one can hypothesize that humans naturally form such sentences, using a mental analog of ‘!’; But
perhaps theorists should posit sparer conceptual capacities in accounts of meaning for human
languages that permit expressions like ‘is above something that was given’.

In terms of explanatory order, one can speculate that humans form and use concepts like
\(\exists z[\text{ABOVE}(X, Y)! \text{GIVE}(Y, Z, W)]\) independent of capacities to generate and understand linguistic
expressions. Then one might appeal to such concepts in theories of linguistic meaning.
Alternatively, one might try to explain linguistic understanding in the sparsest possible way—

\(^{30}\) Using ‘&’ fosters the illusion that the posited species of conjunction is mundane. I assume that human inquirers
can invent and come to think with technical concepts whose properties make them unnatural for purposes of lexical
acquisition and ordinary human thought. Consider \(\text{twix}(\text{p}, \text{q}, \text{r})\): \(\text{p}\) and \(\text{r}\) are equivalent with respect to truth—both
are true, or both are false—while \(\text{p}\) and \(\text{q}\) are not equivalent in this respect. Or \(\text{maj}(\text{p}, \text{q}, \text{r})\): exactly two of the three
thoughts are true. Note that \(\text{twix}(\text{p}, \text{q}, \text{r}) \& \sim(\text{q})\) implies \(\text{maj}(\text{p}, \text{q}, \text{r})\). Compare ‘grue’ (Goodman [1954]), ‘urp’
(Quine [1960] on undetached rabbit parts), and nonconservative determiners (Barwise and Cooper [1981],
Higginbotham and May [1981]).
posing only the concepts and combination operations required in order to account for the relevant natural human phenomena—and argue that this minimalist approach yields a better account of meaning in the following sense: it explains, without sacrificing descriptive adequacy, many facts that are puzzling given richer assumptions about human concepts; cp. Chomsky (1995). Then theorists are free to use the posited concepts and spare combination operations in explaining the unbounded cognitive productivity of human concepts. The latter approach is more ambitious. But I’ll argue that it is also more empirically plausible.

5.3 Minimal Polyadicity (Again)

Ugly facts will, no doubt, force us to a view more complicated than the one urged here. But positing an ideal Fregean/Tarskian mind, to account for human linguistic competence, seems like bad methodology. So I approach the issues from a different direction. Of course, nobody thinks that human heads house ideal minds. But many semantic theories invoke powerful machinery. This makes it tempting for authors to hedge about whether, or in what sense, they are offering empirical claims about human psychology; see, e.g., Davidson (1967b). My preference is to start with an overtly psychological hypothesis that may be too simplistic, and ask what else we need to posit. Instead of assuming that we can form concepts like \( \exists z[\text{AFTER}(x, y) \land \text{GIVE}(y, z, w)] \), and asking how we deploy this capacity in comprehending speech, I assume that we can use linguistic expressions as instructions for how to build concepts like \( \text{BROWN}(\_) \land \text{COW}(\_) \) and \( \exists [\text{AFTER}(\_, \_) \land \text{COW}(\_) \); where this latter concept applies to things above a cow.

Recalling the discussion above, one can characterize M-junction in terms of the concept \( \cap ([\Phi(\_), \Psi(\_)]) \). Though instead of describing ‘cow’ as an expression of type \(<e, t>\)—taking the types \(<e>\) and \(<t>\) as basic for human language semantics—one might describe ‘cow’ as an expression/Begrißfsplan type \(<M>\) for monadic. And one might describe M-junction as an operation of type \(<M, <M, M>>\). One can likewise describe Θ-junction as an operation of type \(<D, <M, M>>\); where ‘D’ indicates a secondary, very limited semantic type. While Θ-junction permits construction of \( \exists [\text{AFTER}(\_, \_) \land \text{COW}(\_) \], the idea is not to generate concepts like \( \exists x[\text{GIVE}(z, y, x) \land \text{COW}(x)] \) or \( \exists x[\text{AFTER}(y, z) \land \text{GIVE}(z, y, x)] \). Even the monadic concept \( \exists x[\text{COW}(y) \land \text{BROWN}(x)] \) fails to be generable via M-junction and/or Θ-junction. And as discussed in chapter five, human languages may not permit instructions of type \(<e>\). In which case, the basic instruction types may be \(<D>\) and \(<M>\), with no non-basic types, since all complex expressions are of type \(<M>\); cp. Partee (2006).

Distinguishing semantic instruction types from Fregean conceptual types, and not appealing to \(<t>\) as a basic instruction type—and certainly not invoking an existential quantifier of type \(<et, t>\), much less a more flexible quantifier that can combine with any predicative concept—highlights the question of what a sentence is. I take this to be a good thing, since as noted above, our best current conceptions of generative grammar make it unclear which if any generable expressions can plausibly be identified with sentences. This unclarity is unsurprising if the notion of sentence has its home in subject-predicate conceptions of thought, while the notion of subject has no stable place in generative grammar; cp. Chomsky (1965). It may be that the only natural sentences of type \(<t>\) belong to languages of thought with roots that are phylogenetically older than the spoken/signed languages that human children acquire. And perhaps expressions of these newer languages are used to build complex monadic concepts, including some special cases that are importantly equivalent to natural thoughts of type \(<t>\), without ever generating expressions of type \(<t>\). In which case, the very idea of a truth conditional semantics for a human language may be fundamentally misguided.
Such speculations aside, Tarski (1933) himself showed how closed sentences could be viewed as special cases of monadic predicates that have satisfaction conditions, and why this does not require appeal to truth values; see chapter two. Drawing on Tarski, I’ll introduce a pair of operations/instructions that correspond to existential closure and negated existential closure of a monadic concept. But these will be of type <M, M>; and <M, M> is not <et, t> in disguise. The idea, spelled out in chapter six, is that \( \text{cow}() \) can be used to build either of two monadic concepts—\( \text{↑cow}() \) or \( \text{↓cow}() \)—that can be used to think about everything or nothing, depending on whether or not there is a cow: if \( \text{cow}() \) applies to something, then \( \text{↑cow}() \) applies to each thing, and \( \text{↓cow}() \) applies to nothing; if \( \text{cow}() \) applies to nothing, then \( \text{↑cow}() \) applies to nothing, and \( \text{↓cow}() \) applies to each thing. The arrows—suggesting upward and downward entailment, modern versions of inference patterns discussed in section two—can only attach to monadic concepts: \( \text{↑cow}() \) is equivalent to though not identical to \( \exists x [\text{cow}(x)] \); by contrast, \( \text{↑above}() \) is neither generable nor coherent.

Propositional conjunction is easily reconstructed in these terms. This in turn invites a proposal about the kind of abstraction required for relative clauses and quantificational expressions like ‘chased every cow that most dogs chased’. One might think it obvious that such expressions are not instructions for how to assemble monadic concepts, even if quantifiers raise and leave traces of displacement, as in the following (partial) syntactic analysis of ‘Fido chased every cow’: \([\text{every cow}, fido [\text{chased} t_1]]\). But chapter six shows how to treat traces as instructions of type \(<M>\), akin to the proper noun in ‘the John we know’, and how to characterize the meaning of ‘every cow’ in terms of \(\Theta\)-junction.

I’ll argue that if quantificational expressions are instructions for how to build concepts of the form \( \exists [\Theta(\_, \_) \wedge \Phi(\_)] \), then these expressions are no harder to accommodate than on views that appeal to saturation and concepts of type \(<et, et, t>\). In fact, the constraints make it easier to isolate what is really required to accommodate quantificational constructions. Similar points apply, I claim, to plural constructions like ‘chased three dogs’. Specialists will want to see the details, including how appeal to truth values can be avoided when it comes to introducing analogs of (nonfirstorderizable) concepts like \( \text{most}[\Phi(\_), \Psi(\_)] \), which cannot be captured in terms of Tarski’s (1933) apparatus. But these are matters for chapters five and six.

5.4 Following Instructions

Let me instead end this first chapter by returning to the relevant notion of instruction and an important sense in which it is intentional. My suggestion is not that ‘gave a cow a dog’ is an instruction to build a concept with which one can think about things that gave a cow a dog. This instruction might be executed by building the concept \( \exists y [\exists z [\text{gave}(x, y, z) \wedge \text{cow}(z) \wedge \text{dog}(y)] \), which has a triadic constituent. My claim is that ‘gave a cow a dog’ is an instruction for how to build a concept like \([\text{give}(\_), \text{past}()^\wedge \exists [\text{recipient}(\_, \_) \wedge \text{cow}(\_)]^\wedge \exists [\text{patient}(\_, \_), \text{dog}(\_)] \), which has only an occasional dyadic concept that is “sealed in.” This M-junction has constituents that include two \(\Theta\)-junctions. For some purposes, the differences between the concepts may not matter. But formal distinctions can be crucial when the issues concern the natural generative procedures that underlie human linguistic competence and cognitive productivity.

Put another way, to satisfy semantic instructions of the sort envisioned here—in the sense in which an instruction can be satisfied by suitable activity, as opposed to an invented sentence being satisfied certain entities—monadic concepts must be assembled in certain ways from building blocks of certain sorts. Assembling monadic concepts in other ways does not satisfy the instructions. (Compare: I told you to make ice cubes from the tap water, not the expensive
bottled water.) So executing assembly instructions may but need not yield concepts that have satisfaction conditions in some other sense.

If meanings are human Begriffsplans (concept assembly instructions) that have mechanical execution conditions, executing these instructions need not lead to ideal concepts. Though in some cases, the assembled concepts may approximate Fregean models. When conditions are optimal, following a recipe may lead to an especially good instance of a product type, even if the recipe is not an instruction for how to create especially good instances of that type. In a kitchen stocked with excellent ingredients, a skilled baker might create a sublime pie by following the procedure specified in a certain book, while another baker in another kitchen makes a merely adequate pie by following that same procedure. Often, an instruction can be carried out in many ways, yielding results that reflect many factors. To take a nonculinary example, suppose we are each told to find a red box and a blue box, and then put the former on the latter. The stability of our stacks will depend on the size and rigidity of the available boxes. But the instruction can be satisfied, even if the result is unstable or otherwise less than ideal.

As we’ll see, if meanings satisfy demanding compositionality constraints, then the corresponding concepts may “fit each other” better than they “fit the world.” For many purposes, it may do no harm to suppose that ∃Y∃Z[GIVE(X, Y, Z) ∨ COW(Z) ∨ DOG(Y)] is “extensionally” equivalent to its natural analog. But this idealization may be limited in ways that matter for both meaning/polysemous in psychology and truth/logic. My suspicion is that most natural concepts do not have extensions; cp. Travis (1985, 200x), Sainsbury (1996). But in any case, I deny that meanings are instructions for how to build concepts that exhibit classical semantic properties, even if executing such instructions sometimes leads to assembly of concepts that have classical semantic properties.  

Procedure matters. In my view, the posited instructions/Begriffsplans have mechanical execution conditions that make no reference to the things we usually think and talk about. Indeed, at least with regard to the core recursive cases, the steps are rather simple-minded: fetch, JOIN, repeat.

From this perspective, theories of human linguistic meaning are theories of how human linguistic expressions are related to human concepts, whose relation to truth may turn out to be quite complicated and orthogonal to the central issues concerning natural semantic composition. This internalistic conception of meaning can be viewed as a variant on the old idea that meanings are concepts—mental representations with which thinkers can think about things—as opposed to things about which speakers can speak. But this runs counter to a newer tradition, according to which meanings (if such there be) are more like referents, and meaning is a lot like referring.

This sets the agenda for chapter two, which describes a possible language of thought, PL, that has a truth conditional semantics: sentences have satisfaction conditions that are recursively specified in terms of a given domain, and truth is characterized in terms of satisfaction; see Tarski (1933). In chapters five and six, I describe another possible language of thought, SMPL. This second-order but massively monadic predicate calculus permits plural predicates, as discussed by Boolos (1998). In some respects, SMPL is more restricted than PL; in other respects, it is richer and more Fregean. That is, SMPL is an extension of a fragment of PL. Eventually, my proposal will be that meanings are instructions for how to build (concepts that

31 In earlier work (e.g., Pietroski 2005b), I simply described meanings as instructions to build concepts, without worrying about the distinction between commands and blueprints: ‘instruction’ is polysemous in a way that lets us talk about an instruction to Φ or an instruction for how to Φ. I had mistakenly thought that analogies to executable programs would make the idea clear enough, and that talk of instructions could be accepted as relatively innocuous without further terminological fuss.
are) predicates of SMPL by fetching atomic predicates and combining them in the simple ways outlined in chapter one. As we’ll see, this account of how meaning is related to psychology has implications for how meaning is (not) related to truth, while the Davidson-Montague conjecture has implications for how meaning is (not) related to psychology. But first, we need to know what it is for any systematically productive language— invented or natural—to have a truth conditional semantics.

Chapters five and six also return to details concerning lexicalization, concept introduction, and semantic composition in human languages. The idea is that concepts like GIVE(＿), introduced so that they can be M/Θ-joined with others, tend to fit the world less well than prior concepts like GIVE(＿, ＿, ＿). Condensing a long line of thought, and leaving out the qualifications: GIVE(＿) does not have mind-independent satisfiers; though unlike UNICORN(＿), GIVE(＿) is lexically related to a concept that does apply to real things; but unlike GIVE(＿), GIVE(＿, ＿, ＿) cannot be systematically combined with other concepts, at least not via the operations invoked by phrasal syntax. Condensing further, neither GIVE(＿) nor GIVE(＿, ＿, ＿) has classical semantic properties: the introduced monadic concept is too far removed from the world; the prior relational concept does not have the right form. As Frege and Tarski knew, inventing a language with a truth conditional semantics is hard enough. It takes work to sync form and content in the requisite ways.
Chapter Two: Invention and Satisfaction

The previous chapter was largely an attempt to explain my claim that phrasal meanings are instructions for how to build conjunctive monadic concepts. This proposal is developed in chapters five and six. Towards the end of chapter zero, I reviewed some reasons—further articulated in chapters three and four—for seeking an alternative to truth-theoretic conceptions of meaning for human languages. This chapter reviews some relevant history and technical notions.

Section one is a reminder that Frege did not provide an algorithm that specified Bedeutungen (or Sinnen) for the boundlessly many expressions of his Begriffsschrift. In this sense, Frege did not provide an explicit semantics for his ideal language; he “merely” offered fecund suggestions for how to interpret his insightful invented syntax. As discussed in sections two and three, Tarski provided a semantics for a first-order fragment of Frege’s language, and Church offered a more Fregean semantics for a more Fregean extension of Tarski’s language. Davidson suggested that we can provide good models of human languages by drawing on Tarski directly, without the Church-style additions, which Montague then added. My suggestion is that we should go back to Frege and Tarski, try to develop a second-order extension of a fragment of a first-order predicate calculus that generates boundlessly many monadic concepts. To understand this dialectic, it helps to first be clear about what Frege didn’t do.

1. You Can’t Always Get What You Want

In chapter one, I described Frege’s taxonomy of concepts in formal terms, setting aside questions about what concepts signify. But Frege wanted expressions of his Begriffsschrift to have a cluster of interpretational properties: each expression of type <e> indicates a particular thing; each expression of type <t> indicates one of two truth values; every other expression indicates a function; and each complex expression, which is the result of combining some expression of a type <α, β> with an expression of type <α>, indicates the result of applying the corresponding function of type <α, β> to the entity or function indicated by the “saturater” of type <α>. Each expression is also supposed to indicate something in a certain way; where this way of indicating (Sinn) determines that which is indicated (Bedeutung), so that a Sinn—a way of thinking about a Bedeutung—can be a constituent of a thought-content (Gedanke) that can be a premise or conclusion of an argument. Frege’s hope was that an ideal thinker would be positioned to know which entity or function is indicated by any ideal expression/concept, given knowledge of which Sinnen the atomic expressions/concepts expressed. But this picture suffered from three serious flaws (see Dummett 1981 and Evans 1982): reference to the unsaturated (§1.1); the persistent specter of Sinn without Bedeutung (§1.2); and lack of an explicit procedure for interpreting Begriffsschrift analogs of ‘Rxy & Syzw’, which can be bound by multiple quantifiers (§1.3).

1.1 Functions and Unsaturated

As noted in the introductory chapter, Frege accorded a kind of priority to the notion of a function as a procedure that maps inputs onto outputs. By itself, that would be fine. But Frege also suggested that functions, in his sense, are unsaturated in a way reflected by ideal predicative symbols. And that lead to headaches.

The basic distinction between procedures and sets is clear enough, at least since Church’s (1941) reconstruction of it. Recall that each “function in intension” (procedure) determines exactly one “function in extension” (set of ordered pairs), which can be determined by many procedures. So we can speak of I-functions vs. E-functions, and say that Frege (1892b) characterized E-functions in terms of I-functions. The E-functions indicated with technical predicates like ‘Number( )’ and ‘Prime( )’ are illustrative cases. In thinking about any such set,
which has infinitely many elements, one thinks about it as the extension of some I-function. So in light of Frege’s (1892a) related Sinn/Bedeutung distinction, one might say that an ideal predicative expression indicates an E-function via an I-function that can itself be thought of in many ways. But Frege did not stop here.

Frege also stressed that thoughts, which can be endorsed and so used to make judgments, are not mere lists of representations. One needs to be careful with the terminology, since words like ‘thought’ and ‘concept’ exhibit symbol/content polysemy; ‘thought’ and ‘judgment’ are also like ‘assertion’, which can be used to describe certain events that be characterized in terms of contents. In speaking of a thought that Sadie is a horse, one might be talking about a mental episode, a mental sentence, or a content shared by various sentences of various languages. But as in chapter one, let’s use small capitals so that HORSE(SADIE) is a mental sentence in which the concept SADIE saturates the unsaturated concept HORSE( ), which is effectively a sentence frame. By contrast, HORSEHOOD is a denoting concept with which one thinks about an abstract entity.

This doesn’t yet require that sentences indicate truth values, or that monadic concepts indicate functions. But we can stipulate that the “ideal” concept HORSE( ) indicates or “Bedeuts” a function from entities to truth values, while SADIE indicates a unique entity in the domain of that function, and the ideal mental sentence HORSE(SADIE) indicates a truth value.¹ The sense (Sinn) of this sentence, which presents the truth value in a certain way, is then a thought-content (Gedanke) that can be endorsed or rejected by endorsing or rejecting HORSE(SADIE). But the predicative concept HORSE( ) is not singular. The indicated Gedanke does not combine a singular way of thinking about Sadie combined with a singular way of thinking about a function. According to Frege, the indicated Gedanke is the result of saturating an unsaturated way of thinking about (things mapped to truth iff they are) horses; where saturating an unsaturated can result in a unified judgeable thought, as opposed to a mere list of constituents.

That does not yet explain the unity in question. But Frege’s hunch was that unsaturated constituents play a special role in explaining the unity of sentences/Gedanken, while saturating constituents play a special role in explaining how expressions/Sinnen can be about things. Part of the idea was that SADIE does not merely indicate its Bedeutung. The saturating concept bears a special relation to a particular horse—a relation that HORSE( ) does not bear to the indicated function. We can also say that SADIE denotes Sadie, while HORSE( ) is of the wrong type to be a denoter; see section four. Frege spoke of HORSE( ) having an unsaturated content. And given his interest in ideal contents, as opposed to actual or even idealized symbols, this led Frege to adopt a notion of ‘Concept’ such that while HORSE( ) sort of indicates a Concept, we should not say that the Concept indicated with HORSE( ) is a Concept: to use ‘the Concept’ and ‘a Concept’ in this way would be to imply that Concepts can be denoted and quantified over.

Frege wanted to refrain from endorsing this implication, and he wanted his ideal language to preclude its formulation; cp. Wittgenstein (1921), Ricketts (199x). But we needn’t dwell on ineffability. Suffice it to say that Frege’s discussion of Concepts seemed at odds with his claim that we can’t talk about them. Even recognizing the motivation for Frege’s claim, and giving him the “pinch of salt” he asked for, one wonders if talk of the unsaturated shouldn’t be tamed and limited to talk of certain symbols.² In any case, until Church helped sort things out, one could

---
¹ We can also define ‘indicates’ so that ordinary mental sentences indicate truth values, even if there is no independent natural sense in which concepts like HORSE( ) indicate functions.
² That leaves the question of what distinguishes a thought from a short list of symbols that exhibit the types <α> and <α, t> for some type <α>; cp. note X chapter two. But one can say that in the thought, one concept saturates the other. That labels the distinction without explaining it, since we don’t yet know concepts are, or what saturation is.
easily get confused by the various Fregean notions of procedure and determination: Sinnen as procedures that determine Bedeutungen; I-functions that determine E-functions; singular concepts that determine/denote entities; and undenotable Concepts that determine functions. As we’ll see, Tarski appealed to none of this, and so avoided the potential confusions.

1.2 Thinking of Nothing

Setting aside reference to unsaturates, concepts like VULCAN and UNICORN( ) suggest a second difficulty for Frege’s intended interpretation of his ideal symbols. Frege was glib about concepts that indicate nothing, as if he assumed that such concepts only figure in episodes of thinking that ideal thinkers would recognize as episodes of story-telling, in a sense of ‘story’ that excludes sufficiently disciplined attempts to think true thoughts about mind-independent aspects of reality. But a serious thinker can mistakenly think that Vulcan is as real as Neptune, that unicorns are odd horses, that phlogiston is a component of air, etc.

To engage with such thoughts, if only to deny them, it seems that one must interpret symbols that lack Bedeutungen. One can say that such symbols still have Sinnen. But as Evans (1982) stresses, if Sinnen are characterized as ways of presenting Bedeutungen, we need some account of “ways of presenting” that applies to examples like VULCAN. This makes it tempting to identify Sinnen with descriptions; see Russell (1905), Quine (1963). But even setting aside Kripke’s (1980) apparently fatal objections, and Frege’s inclination to view descriptions like ‘the smallest prime’ as complex names, one wants to understand conditionals like the following: if there is a greatest prime, then it is greater than its factorial plus one. Perhaps ideal thinkers know that in proofs by reductio, a singular concept is used in a story-telling way. But in the course of rejecting the thought that there is a greatest prime, it seems that even an ideal thinker briefly entertains a thought with a singular component that purports to indicate the greatest prime.

At a minimum, Frege did not say enough about concepts that indicate nothing. Even a great logician can suspect that every predicate (and hence ‘nonselfelemental’) has an extension, and hence that a singular concept can be introduced to denote the set of nonselfelemental sets. This point here is not merely that Frege was susceptible to human error in allowing for such a set. Given paradoxes like Russell’s, one might worry that the interpretive apparatus of functions is unstable—and that this instability is a potential source of massive error that precludes Fregean thoughts from ever being true (without also being not true). Even idealized symbols can fail to make the intended contact with reality, at least if one is not very careful about the kinds of stipulations used to specify interpretations for the symbols. And while Frege’s Begriffsschrift was designed to permit quantification into unsaturated positions (see §1.5), his claims about predicates were inconsistent. By contrast, as discussed below, Tarski characterized his notion of satisfaction for languages that do not permit quantification into predicate positions.

Moreover, in a Tarskian language, only sentences have satisfaction conditions. No interpretations are assigned to subentential expressions. So there is no issue about which interpretation to assign to UNICORN( ) or VULCAN. This makes it easier to stipulate coherent interpretations; though of course, human languages may not conform to these simplifications.

Let me briefly digress, though, to note a different response to concepts without Bedeutungen. One might adopt a hyper-externalistic understanding of ‘ideal’, so that no ideal thinker has the concept VULCAN, and there is no ideal concept VULCAN that lacks a Bedeutung; cp. Evans (1982), McDowell (1995). But then an ideal thinker must not only have her mental house in good order, avoid paradox-inducing claims, and be largely en rapport with an environs free of Cartesian demons. In this hyper-externalistic sense, ideal thinkers are immune from certain kinds of errors, perhaps to the point of being unable to think about the same thing in two
willingly (see, e.g., Heck and May (2006, 2010). Model theory for ideal mental sentences is a bad party game: pretend that thoughts can change contents, like people can change shirts, and ask which contents a thought-form can support given constraints pulled from a hat. Chomsky’s (1965, 1970, 1986, 1995, 2000a) conception of syntax is like Frege’s in this respect, and claims about the “autonomy” of syntax should be understood accordingly (see Berwick et al. 2011); cp. note 16 of chapter zero.

3 See, e.g., Heck and May (2006, 2010). Model theory for ideal mental sentences is a bad party game: pretend that thoughts can change contents, like people can change shirts, and ask which contents a thought-form can support given constraints pulled from a hat. Chomsky’s (1965, 1970, 1986, 1995, 2000a) conception of syntax is like Frege’s in this respect, and claims about the “autonomy” of syntax should be understood accordingly (see Berwick et al. 2011); cp. note 16 of chapter zero.
Since $\exists x \forall y \exists z [\text{ABOVE}(x, y) \land \text{GAVE}(z, w, y)]$ has exactly one unbound variable/slot, it must indicate a function of type $\langle e, t \rangle$. In this case, we know which function is intended: the smallest one that maps an entity (w) to $T$ iff something (x) is above each entity (y) that something (z) gave to the first entity (w). But how is a thinker supposed to work this out, as opposed to just “seeing” which ideal thought is intended? Given a complete thought like $\exists x \forall y \exists z [\text{ABOVE}(x, y) \land \text{GAVE}(z, \text{BESSIE}, y)]$, an ideal thinker could form the thought-frame $\exists x \forall y \exists z [\text{ABOVE}(x, y) \land \text{GAVE}(z, w, y)]$. But the question is how the complete thought is formed in the first place. One can say that $\exists x [\Phi(x)]$—a.k.a. $\exists [\Phi(\_)]$—is a concept of type $\langle et, t \rangle$. But this concept can neither saturate nor be saturated by $\forall y \exists z [\text{ABOVE}(x, y) \land \text{GAVE}(z, w, y)]$. This latter concept is of the dyadic type $\langle e, et \rangle$.

One can introduce quantificational concepts of type $\langle\langle e, et \rangle$, $\langle et, t \rangle\rangle$. But the problem recurs: $\forall y [\Phi(y)]$ is saturable by $\text{COW}(\_)$, but not by $\exists z [\text{ABOVE}(x, y) \land \text{GAVE}(z, w, y)]$, which is of type $\langle e, \langle e, et \rangle \rangle$; and $\exists z [\Phi(z)]$, of the same type as $\exists x [\Phi(x)]$, cannot be saturated by a concept of type $\langle e, \langle e, \langle e, et \rangle \rangle \rangle$. In this sense, Frege helped himself to the powerful idea that $\exists [\Phi(\_)]$ can combine with unsaturated concepts of endlessly many types—in effect, relativizing to assignments of values to any variables not targeted by the quantifier—making his initial suggestion that $\exists [\Phi(\_)]$ is of type $\langle et, t \rangle$ prescient though somewhat naughty.

Again, one can say that ideal thinkers have the requisite abstractive powers, given the capacities required to form thoughts like $\text{ABOVE}($BESSIE, VENUS$)$ $\land \text{GAVE}($SADIE, FIDO, VENUS$)$. But one needs an account of these powers/capacities—which permit of abstraction of a tetradic concept from $\text{ABOVE}($\_, \_) and $\text{GAVE}($\_, \_, \_)—to explain how anyone could form the concepts that Begriffsschrift expressions reflect. Communication raises a related issue: how can a thinker perceive a complex Begriffsschrift expression, whose variables can be bound in various ways by multiple quantifiers, as an instruction to build a correspondingly complex concept? This is not to doubt the utility of Frege’s logical syntax. On the contrary, I take his proposals about the architecture of thoughts to be major contributions to both logic and psychology. But Frege insightfully invented a logical syntax whose intended interpretation raised important questions that he did not answer.

Tarski showed how to interpret formulae of arbitrary adicity, containing variables that can be bound by arbitrarily many existential/universal quantifiers, for certain simple languages. This was another major achievement. But if all human language phrases are instructions to build monadic concepts—or perhaps predicates that need not be viewed as indicating functions of type $\langle e, t \rangle$—theories of meaning for human languages may not require all of Tarski’s apparatus. And if human languages let us build concepts with quantifiers beyond those of the first-order predicate calculus, we may need to think about extensions of fragments of Tarskian languages. That, however, is advertisement. And before turning to Tarski, I want to stress that Frege needed a procedure for interpreting complex quantificational expressions, given how he used his Begriffsschrift. For this is also an opportunity to say how languages can be used to introduce formally new concepts, and how expressions can be Begriffplans (concept assembly instructions), in the context of examples that motivated Frege.

1.4 Thinking of Numbers

The absence of an explicit semantics for his Begriffsschrift was not a mere technical lacuna in the context of Frege’s larger project. Frege maintained that thinkers can employ an ideal language to introduce higher-order concepts that are useful in the study of logic. And especially

---

4 I have been influenced, here and elsewhere, by Hory’s (2007) excellent discussion of Frege’s notion of definition.
when technical concepts are being invented, one cannot assume a capacity to intuit what inventors intend when they use symbols in new ways. For these purposes, an explicit semantics really is called for. (This point is relevant to any version of the idea that human languages can be used to introduce concepts like CHASE(_) or BROWN[_, Φ(_)].) In this context, it will be useful to contrast two kinds of concept introduction: explicit definition of new concepts in terms of old ones; and invention of concepts to be used in theorizing.

Given a concept of predicate conjunction (see §2 of chapter one) and a capacity to introduce concepts by stipulation, one can use HORSE(_), FEMALE(_), and ADULT(_) to introduce a fourth concept: MARE(_) =df ∩[ADULT(_), ∩[FEMALE(_), HORSE(_)]]. From a purely cognitive perspective, it is hard to see the point of such stipulation. But a certain kind of language-user might connect lexical items with atomic concepts. If such a thinker/speaker encounters ‘mare’, she might introduce MARE(_) with a “meaning postulate” that permits replacement of the atomic but not primitive concept with a conjunction. Perhaps BACHLEOR(_) likewise stands in for ∩[¬MARRIED(_), ∩[MARRIAGEABLE(_), MALE(_)]]. Other cases may be more interesting.

Imagine thinkers who use the dyadic concept PARALLEL(_, _) and a concept of existential closure to introduce the triadic concept SHARED-DIRECTION(_, _, _) as follows: 
\[ ∀L∀₁L'∃D[\text{SHARED-DIRECTION}(D, L, L')] \equiv_{\text{df}} \text{PARALLEL}(L, L') \]
This does not determine correctness conditions for all uses of the triadic concept. But for such thinkers, the thought that each line is parallel to itself would imply that any given line shares a direction with itself: \[ ∀L∃D[\text{SHARED-DIRECTION}(D, L, L)] \]. Such introduction raises questions about the proper treatment of variables in discourse that aims at truth. But quantifying over directions given lines seems less ontologically inflationary than quantifying over pencils given lines; cp. Wright (200x). Perhaps properly understood, \[ ∃L∃D[\text{SHARED-DIRECTION}(D, L, L)] \] does not imply the existence of a shared direction any more, or less, than \[ ∃L[\text{PARALLEL}(L, L)] \] does.

In any case, theorists often introduce concepts like QUARK(_) in a more ambitious way that is not definitional. Invented concepts can be related, in various ways, to natural concepts. One might regard non-Euclidean ways of thinking about space, or relativistic ways of thinking about mass, as abstract analogs of more familiar concepts—treating generalizations previously held to be universal as special or limiting cases. Though if the world is fundamentally bizarre, we may need to invent bizarre concepts to think about the world as it is fundamentally. This is not the kind of concept introduction Frege discussed. But for some technical concepts, one can imagine minds that start with them, and then try to characterize our “more primitive” concepts as special case of theirs. Frege undertook such a project with regard to a cluster of arithmetic concepts including NUMBER(_), ZERO, and PRECEDES(_, _).

We can use these concepts to form thoughts that might be regarded as axiomatic: zero is a number; every number precedes another; no number precedes zero; etc. But famously, Frege wanted to encode such thoughts in a way that captures what follows from them, while also making it possible to ask what they follow from. He hoped to reduce (his versions of) the Dedekind-Peano axioms for arithmetic to principles of logic and definitions for the requisite nonlogical vocabulary. He did not fully succeed. But given the definitions he introduced, and using only a consistent fragment of his logic, Frege was able to derive the (re-presented) axioms from “Hume’s Principle” concerning the relation of cardinality to correspondence:

\[ (HP) \forall s \forall s' \{ \equiv [-^c, c^c \text{NUMBER-OF}(S), ^c^c \text{NUMBER-OF}(S')] \equiv_{\text{df}} \text{ONE-TO-ONE}(S, S') \}\]

\[ ^{5}\text{Cp. Frege (1884). The triadic concept might then be used to introduce concepts of having a direction and directions: DIRECTION(D) =_{\text{df}} ∃L[\text{DIRECTION-OF}(D, L)]; DIRECTION-OF(D, L) =_{\text{df}} ∃L'[\text{SHARED-DIRECTION}(D, L, L')]. }\]
two sets have the same cardinality iff their elements correspond one to one.\(^6\)

As the superscript suggests, \(<e, e> \text{NUMBER-OF}()\) is not a concept of type \(<e, t>\); though \(\text{NUMBER-OF}(, \_)\) is of type \(<e, et>\). Saturating \(<e, e> \text{NUMBER-OF}()\) with an \(<e>\)-type concept that denotes a set yields a complex \(<e>\)-type concept of the number that is the set's cardinality. The concepts \(=_[, \_]\) and \(\text{ONE-TO-ONE}(, \_)\) are of type \(<e, et>\), taking numbers and sets to be entities, while \(\equiv_[, \_]\) is of type \(<t, tt>\). So (HP), as encoded above, reflects a relation exhibited by these three equivalence relations. But if the numbers are to be characterized in terms of logic and (HP), as opposed to logic and the unreduced Dedekind-Peano axioms, reconceptualization is required.

In particular, one has to imagine a line of thought in which the concept \(\text{NUMBER}(\_)\)—which applies to things like zero and thirteen—is introduced, as opposed to being a primitive concept. But that's not so surprising if you think about arithmetic induction the way Frege did. It seems obvious that for any property \(P\): if (i) zero has \(P\), and (ii) any given number has \(P\) if its predecessor does, then (iii) every number has \(P\). Yet replacing ‘number’ with ‘cow’, in (ii) and (iii), does not preserve obviousness. The contents of \(\text{NUMBER}(\_), \text{ZERO},\) and \(\text{PREDECESSOR}(, \_)\) are evidently related in a way that the contents of \(\text{COW}(\_), \text{ZERO},\) and \(\text{PREDECESSOR}(, \_)\) are not. This invites the project of re-presenting the contents of arithmetic concepts in more interesting ways, and perhaps *deriving* a reformulated principle of arithmetic induction.

To this end, Frege defined zero as the number of (the set of) nonselfidentical things, and defined the other numbers inductively. One, whose predecessor is zero, is the number of things identical with zero; two, whose predecessor is one, is the number of things identical with zero or one; etc. Each number \(n\) is the predecessor of another—viz., the number of numbers in the “predecessor series” up to and including \(n\); and given (HP), \(n\) is the number of set \(s\) iff the elements of \(s\) correspond one to one with the numbers in the predecessor series up to but not including \(n\). By itself, this is merely a clever characterization of the numbers. But Frege also introduced \(\text{ANCESTRAL-R}(, \_)\)\(^7\)—a higher-order concept of type \(<e, et>, <e, et>\)\(^>{}\)—which can be saturated by a relational concept like \(\text{PREDECESSOR}(, \_)\) to yield a concept whose extension is the transitive closure of the saturating concept. Indeed, this is where Frege really put his relational logic to work. Given any type \(<t>\), one can introduce expressions of the higher type \(<\tau, \tau>\). In particular, \(\text{ANCESTRAL-R}(, \_)\) can be defined shown below.

\[
\forall x \forall y \{\text{ANCESTRAL-R}(x, y) \equiv_{df} \exists z [R(x, z) \& \forall f [F(x) \& \forall w \forall z [F(w) \& R(w, z) \supset F(z)] \supset F(y)]]\}.
\]

That is, \(x\) is an \(R\)-ancestor of \(y\) iff \(x\) bears relation \(R\) to something, and every function \(F\) of type \(<e, t>\) meets the following condition: if \(F\) maps \(x\) to \(t\), and \(F\) maps \(z\) to \(t\) whenever anything that \(F\) maps to \(t\) is \(R\)-related to \(z\), then \(F\) maps \(y\) to \(t\).\(^7\)

In general, for any relation and any entities: \(\text{ANCESTRAL-R}(x, z)\) if \(R(x, y)\) and \(R(y, z)\).

The relation of precedence, which zero bears to two, is the ancestral of the predecessor relation that zero bears to one and no other number: \(\forall x \forall y [\text{PRECEDES}(x, y) \equiv \text{ANCESTRAL-PREDECESSOR}(x, y)]\). This makes the logical relation between \(\text{PRECEDES}(, \_)\) and \(\text{PREDECESSOR}(, \_)\) manifest;

---

6 See Wright (198x), Heck (2011), Demopolous (1993), Zalta (200x). Frege’s version of (HP) quantified over “courses of values” of functions, as opposed to sets per se. And in some settings, one might want to relativize to idealized concepts. For as Frege notes, thousands of soldiers can be a few regiments and a single army.

7 Put another way, \(x\) is an \(R\)-ancestor of \(y\) iff \(x\) bears \(R\) to something, and whatever the Fs may be: \(y\) is one of them *if* \(x\) is one of them, and an entity \(z\) is one of them whenever one of them is \(R\)-related to \(z\). The idea is that \(x\) is an ancestor of \(y\) if \(y\) has every “hereditary property” that \(x\) “passed on;” where an individual, Eve, may pass on properties like *being human* or *being descended from* or *identical with* *Eve.*
compare ANCESTOR(, _) and PARENT(, _). One can also say that \( n \) is in the “predecessor series” up to and including \( m \) iff \( n \) precedes—i.e., is ancestral predecessor of—or is identical to \( m \).

By representing the contents of \( \text{PRECEDES}(, _) \) and \( \text{ZERO} \) in terms of logical notions and concepts used in (HP), Frege licenced replacement of \( \text{PRECEDES}(, _) \) and \( \text{ZERO} \) with complex concepts that manifest logical relations to \( \preceq \). \( \text{NUMBER-OF}(_) \). He could then define the numbers as those things preceded by, or identical to, the number of nonselfidentical things.

\[ \text{NUMBER}(_) \equiv \text{OR}[(\text{ZERO},_), \text{PRECEDES}(_{\text{ZERO}})] \]

This showed how a mind furnished with the concepts needed to formulate (HP) could use certain logical capacities—including a capacity to introduce new concepts via logical notions like \( \text{ANCESTRAL-R}(, _) \)—to reconstruct our concept \( \text{NUMBER}(_) \), along with related arithmetic concepts, in a way that would permit agreement on the Dedekind-Peano generalizations.

Initially, we might view \( \preceq \) as mental shorthand: \( n \) is the number of \( s \) iff \( n \) is a number such that the numbers preceding \( n \) correspond one-to-one with the elements of \( s \).

(Take of counting the elements of a set, starting with zero.) This takes our concept \( \text{NUMBER}(_) \), or perhaps the corresponding Aristotelian predicate, as given. But for Frege, the point is not to reduce \( \preceq \) to more familiar notions, or to argue that we always had the structured concept \( \text{OR}[(\text{ZERO},_), \text{PRECEDES}(_{\text{ZERO}})] \). The idea is rather that given certain “fruitful” concepts, which we may initially find unfamiliar, an ideal thinker could do arithmetic by introducing \( \text{NUMBER}(_) \) as an atomic but not primitive monadic concept.

One can say that for us, as opposed to ideal thinkers who might think about our concepts, Frege’s definitions are not analytic. Still, his definitions reflect a serious attempt to say what numbers are—and how we can think of them—in a way that reflects our capacity to appreciate the compellingness of certain inductive inferences. The idealized lines of thought are available to humans, although introducing the requisite concepts takes work and ingenuity. One must, of course, entertain a thought—e.g. that every number precedes another—before offering a Fregean analysis of its content. But we can use concepts like \( \text{NUMBER}(_) \) and \( \text{PRECEDES}(, _) \) to think, and then introduce technical analogs of these concepts, with the aim of re-presenting our initial thought contents in ways that let us describe certain inferences as valid. Then we can replace \( \text{NUMBER}(_) \) with \( \text{OR}[(\text{ZERO},_), \text{PRECEDES}(_{\text{ZERO}})] \)—or a more structured concept—when this is useful in a proof, and go back to \( \text{NUMBER}(_) \) when the further structure is irrelevant.

Given a fine-grained notion of content, or thought-equivalence, analysis may be impossible or paradoxical. But Frege employed at least two notions of content: one based on his notion of sense (Sinn), and another according to which thoughts are equivalent if each follows from the other. Given latter notion of logical equivalence, or Lewis’ (1975, 198x) characterization of contents as sets of logically possible worlds, one can say that our extant ways of representing are not yet perspicuous. We can use our current concepts to ask questions that lead us to reformulate the questions in ways that allow for interesting answers. Correlatively, analysis can be a creative activity that need not reflect the structure of the representations we initially use to entertain analyses; see Hory (2007) for further discussion.

As I have been stressing, Frege introduced higher-order polyadic analogs of monadic concepts. In this respect, my project is the converse of his. Frege invented logically interesting

---

8 Much less \( \text{OR}[(\preceq, \preceq) \text{NUMBER-OF}((X:=(X, X)),_), \text{ANCESTRAL-PREDECESSOR}(\preceq, \preceq) \text{NUMBER-OF}((X:=(X, X)),_)] \).

9 One can stipulate that definitions, if such there be, meet demanding conditions not met by dictionary entries or Frege’s descriptions of numbers. But I suspect that Frege (like Aristotle) was more interested in defining things—as opposed to words—and refining pre-theoretic concepts, in order to better track theoretically important aspects of the things we think about.
concepts, and treated monadicity as a special case of relationality to truth (see §4.2), as part of a project in logic that prescinds from many details of human psychology. I think humans naturally use concepts of various adicities to introduce logically boring monadic analogs. But I adopt Frege’s idea that extant concepts can be used to introduce formally new ones, and that introducing type-distinct analogs of extant concepts can be useful for certain derivational purposes. Frege “unpacked” monadic concepts like NUMBER(_), in ways that let him exploit the power of his sophisticated polyadic logic to derive arithmetic axioms from (HP). I am suggesting that human languages let us ignore adicity distinctions like those exhibited by CHASE(_ , ) and GIVE(_, _, _)—in favor of monadic concepts like CHASE(_ ) and GIVE(_), which can be combined in simple ways that allow for simple inferences like conjunction reduction. But the big idea, which I am applying to the study of human languages in a certain way, is from Frege: languages are not mere tools for expressing available concepts; they can be used to introduce formally new concepts that are useful given certain computational capacities and limitations. This is the main reason I have dwelt so long on Frege’s project. For while the idea of concept introduction was important for Frege, it is not the part of Frege’s work that semanticists typically draw on.

1.5 Up the Hierarchy
It is also worth noting that Frege’s project requires that we mortals be able to understand notation like ‘ANCESTRAL-PREDECESSOR(_, _)’. This does require an explicit semantics, of the sort provided by Tarski and Church, that specifies the intended interpretation of the initially alien formalism. Concepts like ANCESTRAL-PREDECESSOR(<c NUMBER-OF( {x: x ≠ x} ), _ ) are supposed to be representations with which an ideal thinker—one who does not define these concepts in terms of our logically inferior concepts—can think. So we need a clear characterization of what these representations represent. Put another way, the Bedeutungen for atomic ideal concepts must be specified objectively, in terms of concept-independent entities; and the Bedeutungen for complex ideal concepts must be specified recursively, in a way that human thinkers can follow.

Frege came close to providing enough. He had, as we might now put it, the idea of a truth conditional semantics for an ideal language: a procedure for interpreting, or at least specifying the Bedeutungen of, expressions of a possible language of idealized human thought; cp. Wittgenstein (1921). But it was not unproblematic to presuppose unsaturated concepts that indicate functions which map entities (indicated by saturating concepts) onto truth values, especially given conjunctions and negations of polyadic concepts in the scope of multiple quantifiers. So a stipulated semantics for his invented syntax can be valuable without being a model for human languages. It is one thing to note that logicians need to say what their formalism means, in terms of entities that are independent of our natural concepts, in order to describe the sparsest bundle of concepts that an ideal mind could start with and still construct “logical analogs” of ours. It is quite another thing to hypothesize that children acquire languages that have a semantics in the sense that Frege envisioned; and this was no part of Frege’s view.

A human language may not have an objective semantics in this sense. History suggests that it takes lots of insight and intellectual work to invent formalism that lets us represent some truth conditions systematically. I think it would be stunning to discover that children naturally acquire languages that are already suitable for representing the truth conditions Frege wanted to get at, as if the only real value of his Begriffsschrift is that it was better than German for purposes of conducting proofs. But in any case, even if each Begriffsschrift sentence specifies a truth condition that is shared with some expression of a human I-language, it is worth being explicit about the hierarchy of atomic expression-types that Frege’s invented language permits. This highlights both (i) the importance of specifying interpretations for invented expressions that
are not instances of the basic types \(<e>\) or \(<t>\), and (ii) a respect in which the space of potential atomic Fregean expressions seems to be vastly larger than the space of potential lexical items of a human I-language.

As already noted, expressions of type \(<\tau, \tau>\) can be used to introduce expressions of type \(<\tau, t>\). But this is just a special case of the point that given expressions of types \(<\alpha>\) and \(<\beta>\), one can introduce expressions of type \(<\alpha, \beta>\). So there are many Fregean types. For simplicity, let’s assume exactly two truth values, \(T\) and \(\bot\); where these may be entities of type \(<e>\), perhaps one and zero, chosen at random. The two basic types, \(<e>\) and \(<t>\), can be viewed as constituting the basic “level”—Level Zero—of a hierarchy whose “next” level includes four types: \(<e, e>\); \(<e, t, e>\); \(<t, e>\); and \(<t, t>\); where each of these types corresponds to a class of functions from entities of some Level Zero sort (numbers or truth values) to entities of some Level Zero sort.

Put another way, Level Zero is exhausted by the two types of kind \(<0, 0>\), and Level One is exhausted by the four types of kind \(<0, 0, 0>\). At the next level of this hierarchy, Level Two, there are the new types that can be formed from those at the lower levels (Zero and One):
- eight \(<0, 1>\) types, including \(<e, e, t>>\) and \(<t, t, e>>\);
- eight \(<1, 0>\) types, including \(<e, e, e>>\) and \(<t, t, t>>\); and
- sixteen \(<1, 1>\) types, including \(<e, e, e>>\), \(<e, e, t>>\) and \(<t, t, t>>\).

So at Level Two, there are thirty-two types, each corresponding to a class of functions.\(^{10}\) At the next level, Level Three, there are the 1408 new types that can be formed from those at the lower levels (Zero, One, and Two):
- sixty-four \(<0, 2>\) types, including \(<e, e, e, t>>>\);
- sixty-four \(<2, 0>\) types, including \(<e, e, e, t>>\), \(t>>\);
- one-hundred-and-twenty-eight \(<1, 2>\) types, including \(<e, e, t>, e, t>, e, t>>>\);
- one-hundred-and-twenty-eight \(<2, 1>\) types, including \(<e, e, t>, t, e>>>\); and
- one-thousand-and-twenty-four \(<2, 2>\) types, including \(<e, e, e, e, e, e, e, t>>>\).

At Level Four, there are more than two million types: \(<e, e, e, e, e, t>>>\) and 5631 more examples of type \(<0, 3>\) or \(<3, 0>\); 11,264 of type \(<1, 3>\) or \(<3, 1>\); 90,112 of type \(<2, 3>\) or \(<3, 2>\); and 1,982,464 of type \(<3, 3>\). Let’s not worry about Level Five, at which there are more than \(5 \times 10^{12}\) types. It is, however, worth thinking about Levels Three and Four.

As noted above, Frege’s concept of ancestral is of type \(<e, e, e, t>>>\). This type is at Level Three of the hierarchy, along with more than 1400 others. For each of these types, one can introduce boundlessly many concepts of that type, just as one can introduce \(ANCESTRAL-R(_-, _)\) and boundlessly many other concepts of type \(<e, e, e, t>>>\). One can imagine minds that think in Fregean terms, find his encoding of arithmetic perfectly natural, are able to introduce concepts of any Level Three type, but suffer from performance limitations (e.g., concerning memory) that keep them from using concepts at Level Four or higher. Such thinkers might appreciate that given concepts of type \(<e, e, e, t>>>\), there are concepts of the Level Four type \(<e, e, e, e, e, t>>>\). But this appreciation might be purely theoretical.

One can similarly hypothesize that while human I-languages generate expressions of every type \(<\alpha, \beta>\) such that \(<\alpha>\) and \(<\beta>\) are types, performance limitations keep us from introducing lexical items of types that go beyond those exhibited in human languages. Even if we can appreciate that induction requires going up the hierarchy in the way that Frege describes, grasping this theoretical point may not allow us to introduce concepts of the Level Three type \(<e, e, e, t>, e, e, t>>>\)—even if we can use certain quantifiers and verbs to express concepts of the

\(^{10}\) Compare iterative conceptions of the Zermelo-Frankl sets; see e.g. Boolos (1998).
Level Three types \(<<e, t>, <<e, t>, t>>\) and \(<e, <e, et>>\). But this raises the question of how to describe the alleged performance limitations, along with the types we do use, in a way that yields a plausible account of why certain types are beyond us linguistically though not conceptually; cp. Chomsky’s (1965) discussion of “explanatory adequacy.” My suggestion will be that we get a better overall account by imposing a severe limitation on the space of available types, instead of starting with the entire Fregean hierarchy and trying to filter out boundlessly many types.

2. Tarski’s Satisfaction

There are many ways of presenting the essentials of Tarski’s (1933) semantics for a first-order predicate calculus. Since my aims concern human languages and thought, as opposed to logic per se, it will be useful to describe a possible language of thought, PL, that has a Tarskian semantics; cp. Field (197x). And let’s think of PL as a procedure (an I-language), which generates endlessly many mental sentences from finitely many elements, though without assuming that any actual thinker (apart from Tarski) ever implemented this idealized procedure without special training. As we’ll see, endlessly many sentences of PL are neither true nor false, and the semantics for PL does not specify denotations/Bedeutungen. Still, PL has an objective semantics. Sentences of PL have a kind of significance that is recursively characterized in terms of satisfaction by sequences of entities. And while this significance is not compositional in the way that concepts are, we can define truth for PL in terms of satisfaction. Given this model language, we can go on to ask which additions/subtractions yield better models of human languages.

2.1 A Simple Mentalese

The elements of PL include a variable ‘x’, a variable-extender ‘’’, and finitely many unsaturated concepts of the sort discussed in chapter one: COW( ) and PLANET( ); ABOVE( , ) and CHASE( , ); GIVE( , , ) and BETWEEN( , , ); etc. If n is a variable, so is the result of affixing ‘’’ to n. More briefly, and without quote marks: if n is a variable, so is n’. Hence, there are endlessly many variables of PL: x, x’, x”, etc. For each unsaturated concept, the result of replacing each of its slots with a variable is a sentence of PL. So endlessly many generable sentences—including COW(x), COW(x’), etc.—have no sentential constituents. As discussed in §2.2 of chapter one, a mind might be able to form ABOVE(BESSIE, VENUS) without being able to form ABOVE(BESSIE, ). But we are now imagining a mind that can use the concept ABOVE( , ) to form mental sentences like ABOVE(x’, x) and ABOVE(x, x’), along with ABOVE(x, x) and endlessly many others that employ different variables.

Each unsaturated element (or thought frame) of PL is thus the progenitor of an unbounded family of sentences that have no sentential constituents. But the elemental concepts are not themselves sentences of PL. One can think of basic sentence-formation as a process in which variables are inserted into all the slots of a unsaturated concept. So while GIVE( , ) and GIVE(x, , x’) are not sentences of PL, GIVE(x”, x’, x) and GIVE(x”, x’”, x’) are. The elements of PL also include concepts like BESSIE and VENUS. And for each sentence S of PL, the result of replacing an occurrence of a variable in S with a denoter is another sentence. So the sentences of PL include: COW(BESSIE); ABOVE(x, BESSIE) and ABOVE(VENUS, x); ABOVE(VENUS, BESSIE); etc. We can thus retain the idea that unsaturated concepts can be converted into simple mental sentences by filling slots with singular concepts. But GIVE(x, , BESSIE) is not a sentence of PL. And unlike saturation, merely inserting a variable does not alter adicity: ABOVE(x’, x) is as dyadic as ABOVE( , ).

The sentences ABOVE(VENUS, x) and ABOVE(x, BESSIE) are analogs of the partially saturated thought-frames ABOVE(VENUS, ) and ABOVE( , BESSIE). But the sentences share a variable that reappears in GIVE(x’, x, x’). By contrast, ABOVE(VENUS, ) and ABOVE( , BESSIE)
and $\text{give}(\text{VENUS}, _, \text{BESSIE})$ are merely similar in having an unsaturated slot. Even if a mind can form these complex monadic concepts, it might be unable to link them. We can—and just did—imagine a process of modifying mere unsaturated slots, so that they admit variables as well as saturating concepts, and then adding a device that generates variables. But since mere slots are not yet variables, we should distinguish $\text{above}(\text{VENUS}, _) \equiv \text{above}(\text{VENUS}, X)$.\textsuperscript{11}

The complex sentences of PL include negations and conjunctions: if $S$ is a sentence, so is $\neg S$; if $S$ and $S^*$ are sentences, so is $[S \& S^*]$. Here, I use the usual ampersand, not ‘$\&$’ as in chapter one. As a generable sentence, $[\text{cow}(X) \& \neg \text{above}(X', X')]$ is unobjectionable. It is easy to see how combining two formal objects can create a third that has more variables than either input. Likewise, combining a sentence that has one variable with a sentence that has two variables can yield a sentence that has two variables, one of which has two occurrences as in $[\text{cow}(X) \& \text{above}(X, X')]$. The trick is to show how instances of $\neg S$ and $[S \& S^*]$ can have the intended negative/conjunctive interpretations, no matter how many variables are included, in a way that will also allow for quantificational sentences like $\exists x[\text{cow}(X) \& \neg \text{above}(X', X')]$.

In chapter one, a recurrent question was how complex concepts like $\text{above}(\text{VENUS}, _) \equiv \text{above}(\text{VENUS}, X)$ can be assembled from atomic concepts. If $\text{above}(\text{VENUS}, _) \equiv \text{above}(\text{VENUS}, X)$ is the result of extracting a denoter from a thought, one wants to know why the result is a concept as opposed to formal junk. Variable insertion is part of a potential answer: $\text{above}(_, _) \equiv \text{above}(X', X)$ can be used to build sentences like $\text{above}(X', X)$. But one cannot form a dyadic concept by putting a dime into the first slot of $\text{above}(_, _)$, a penny into its second slot, and then combining the result with the result of inserting another dime into the slot of $\text{cow}(_)$. We can stipulate that variables differ from coins and many other things in this respect. Though even if we assume that $[\text{above}(X', X) \& \text{cow}(X)]$ is a dyadic sentence, one wants to know what it signifies, and why it can apply to pairs of things.

### 2.2 Interpreting Complexity via Sequences

Tarski offered an inspired strategy for assigning interpretations to the sentences of a language like PL. As with the initial discussion of Frege’s typology for concepts, I want to go slowly, since it will be important to distinguish two uses of Tarski’s notion of satisfaction: providing an objective semantics for an invented language, and thereby characterizing a notion of truth that is important for logic as a normative enterprise; and offering empirical hypotheses about human concepts, words, and modes of composition. (For readers who want to skip or skim this review, let me note that the last two paragraphs of this subsection are important for what follows.)

The leading idea is that the variables of PL range over entities that can be ordered as sequences that can satisfy sentences. This does not require that elements of PL denote entities or functions. But it does require entities that can be ordered in a way that corresponds to the generation of variables. As an obvious illustration, if we want to use PL as an invented language for reasoning about sets, we might stipulate a domain of pure sets: $\emptyset$, $\{\emptyset\}$, $\{\{\emptyset\}\}$, etc. If we are studying animals who seem to use PL as a natural mentalese, the relevant domain is less clear. Is

\textsuperscript{11}If any slot can be filled by any variable, and any two sentences can be conjoined, this yields a strong form of systematic composability; cp. Evans (1982), Fodor and Pylyshyn (1988). Given “universal” variables that can be freely replaced with singular concepts, each potential saturater of $\Phi(\_)$ is a potential saturater of $S(\_)$—even if $\Phi(\_)$ and $S(\_)$ originally required saturaters of different sorts, say because these concepts are rooted in different cognitive modules. But a thinker might have various kinds of variables for various kinds of concepts, and some concepts may not interface with just any variable-inserter, leaving pockets of concepts that are not integrated with others. So one might imagine variants of PL that allow for diverse sentences like $\text{sound}(\_)$, $\text{surface}(\_)$, $\text{cause}(\_\', \_\alpha)$, and $\text{between}(\_\", \_\', \_\alpha)$. I’ll return to these points. As suggested in chapter one, I think human concepts are as integrated as they are in part because humans enjoy special linguistic capacities; cp. Spelke (2002, 2003), Carruthers (2002).
it the entire world, as a totality of things? Does the domain include all and only the entities countenanced by correct scientific theories? Does it include cabbages, kings, kingdoms, and things that are “commonsense” entities for humans? Is the relevant domain limited to things, or perhaps concept-independent things, about which the animals under study can think? For now, let’s set these questions aside, and take the notion of a domain entity (henceforth, d-entity) to be clear enough for purposes of characterizing the notion of a sequence.

A sequence is an ordering of the d-entities. The variables of PL are ordered like natural numbers: \( X, X', X'' \), etc. So we can think of sequences as mappings from variables to d-entities. For each sequence \( \sigma \) and any variable \( n \), \( \sigma(n) \) is a d-entity: \( \sigma(x) \) is the first d-entity of \( \sigma \), \( \sigma(x') \) is the second, etc. Since there are boundlessly many variables of PL, even if there are only finitely many d-entities, each d-entity can appear boundlessly many times in each sequence. Crucially, sequences are also plentiful in a technical sense. For each sequence \( \sigma \), variable \( n \), and d-entity \( \Delta \), there is a sequence \( \sigma^* \) such that: \( \sigma^*(n) = \Delta \), and for every other variable \( m \), \( \sigma^*(m) = \sigma(m) \).

Consider a particular sequence, \( \text{BVS} \sigma \), whose first three positions are occupied by (respectively) Bessie, Venus, and Sadie. This sequence, along with endlessly many others, is of the form \( \langle \text{Bessie, Venus, Sadie, ...} \rangle \). Assume that Bessie is a cow, Venus is not a cow, and Venus is above Bessie. Then \( \text{BVS} \sigma(x) \) is a cow, and \( \text{BVS} \sigma(x') \) is above \( \text{BVS} \sigma(x) \). However, \( \text{BVS} \sigma(x') \) is not a cow, and \( \text{BVS} \sigma(x) \) is not above \( \text{BVS} \sigma(x') \). Now consider the sequence \( \text{VBS} \sigma \) that is just like \( \text{BVS} \sigma \) except that first two entities are reversed: \( \text{VBS} \sigma(x) = \text{BVS} \sigma(x') = \text{Venus} \); and \( \text{VBS} \sigma(x') = \text{BVS} \sigma(x) = \text{Bessie} \). Then \( \text{VBS} \sigma \) is of the form \( \langle \text{Venus, Bessie, Sadie, ...} \rangle \). So \( \text{VBS} \sigma(x) \) is not a cow, and \( \text{VBS} \sigma(x') \) is not above \( \text{VBS} \sigma(x) \); but \( \text{VBS} \sigma(x) \) is a cow, and \( \text{VBS} \sigma(x') \) is above \( \text{VBS} \sigma(x') \). If Venus is also above Sadie, then \( \text{VBS} \sigma(x) \) is above \( \text{VBS} \sigma(x'') \), and \( \text{VBS} \sigma(x') \) is above \( \text{VBS} \sigma(x'') \).

Given the notion of a sequence, we can say that many sentences of PL are satisfied by many sequences. In particular, we can say that for any variable \( n \), a sequence \( \sigma \) satisfies \( \text{cows}(n) \) iff \( \text{cows}() \) applies to \( \sigma(n) \). Though for simplicity, let’s just say that \( \sigma \) satisfies \( \text{cows}(n) \) iff \( \sigma(n) \) is a cow. Then \( \text{BVS} \sigma \) satisfies \( \text{cows}(x) \) iff \( \text{BVS} \sigma(x) \) —a.k.a. Bessie—is a cow. So \( \text{BVS} \sigma \) satisfies \( \text{cows}(x) \). But this sentence of PL is not satisfied by \( \text{VBS} \sigma \), since \( \text{VBS} \sigma(x) \) —a.k.a. Venus—is not a cow. Though \( \text{VBS} \sigma \) does satisfy \( \text{cows}(x') \), since \( \text{VBS} \sigma(x') \) is Bessie.

Similarly, let’s say that for any variables \( n \) and \( m \), \( \sigma \) satisfies \( \text{above}(n, m) \) iff \( \sigma(n) \) is above \( \sigma(m) \). Then \( \text{BVS} \sigma \) satisfies \( \text{above}(x', x) \) iff \( \text{BVS} \sigma(x') \) is above \( \text{BVS} \sigma(x) \). Since Venus is above Bessie, \( \text{BVS} \sigma \) satisfies \( \text{above}(x', x) \); but this sentence of PL is not satisfied by \( \text{VBS} \sigma \). By contrast, \( \text{VBS} \sigma \) satisfies the distinct sentence \( \text{above}(x', x) \); though \( \text{BVS} \sigma \) does not satisfy \( \text{above}(x, x') \). Just as \( \text{cows}(x) \) differs from \( \text{cows}(x') \), \( \text{above}(x', x) \) differs from \( \text{above}(x, x') \).

Likewise, \( \langle \text{cows}(x) \& \text{above}(x', x) \rangle \) differs from both \( \langle \text{cows}(x') \& \text{above}(x', x) \rangle \) and \( \langle \text{cows}(x) \& \text{above}(x, x') \rangle \). This makes it easy to specify a sense in which instances of \( \text{cows} & \text{cows}^* \) are conjunctions: a sequence \( \sigma \) satisfies a sentence of this form iff \( \sigma \) satisfies both \( \text{cows} \) and \( \text{cows}^* \). So in particular, \( \text{BVS} \sigma \) satisfies \( \langle \text{cows}(x) \& \text{above}(x', x) \rangle \), since \( \text{BVS} \sigma(x) \) is a cow and \( \text{BVS} \sigma(x') \) is above \( \text{BVS} \sigma(x) \). But neither \( \text{BVS} \sigma \) nor \( \text{VBS} \sigma \) satisfies \( \langle \text{cows}(x') \& \text{above}(x', x) \rangle \), since \( \text{BVS} \sigma(x') \) is not a cow, and \( \text{VBS} \sigma(x') \) is not above \( \text{VBS} \sigma(x) \). By contrast, \( \text{VBS} \sigma \) satisfies \( \langle \text{cows}(x') \& \text{above}(x, x') \rangle \).

Dealing with negation and sentences of arbitrary adicity is also easy. A sentence of the form \( \neg \text{cows} \) is satisfied by \( \sigma \) iff it is not the case that \( \sigma \) satisfies \( \text{cows} \). For any variables \( n, m, k \), \( \sigma \) satisfies \( \text{gave}(n, m, k) \) iff \( \sigma(n) \) gave \( \sigma(m) \) to \( \sigma(k) \). So \( \text{VBS} \sigma \) satisfies \( \neg \langle \text{gave}(x', x', x'') \& \text{cows}(x') \rangle \).

\[ \text{Since sequences are plentiful,} \quad \text{BVS} \sigma \text{ ensures a sequence just like it, except that Venus is the first element. And that sequence,} \quad \text{VBS} \sigma, \text{ ensures a sequence just like it, except that Bessie is the second element.} \]
iff it is not the case that $VBS\sigma$ satisfies both $GAVE(x, x', x'')$ and $COW(x')$. That is, $VBS\sigma$ satisfies the complex sentence iff it is not the case that: Venus gave Bessie to Sadie, and Bessie is a cow. Likewise, $VBS\sigma$ satisfies $[COW(x'') \& ABOVE(x', x)]$ iff Sadie is a cow and Bessie is above Venus. So $VBS\sigma$ does not satisfy this complex sentence, but it does satisfy $\neg[COW(x'') \& ABOVE(x', x)]$.

For any variable $n$ and sequence $\sigma$, we can say that $\sigma(n)$ is the value of $n$ relative to $\sigma$. And if there are no “empty” singular concepts like $VULCAN$, we can say that each singular element of $PL$ is a degenerate variable—a constant—which has the same value relative to each sequence. For each sequence $\sigma$: $\sigma(\text{BESSIE}) = \text{Bessie}; \sigma(\text{VENUS}) = \text{Venus};$ etc. It follows that some sequence satisfies $ABOVE(\text{VENUS}, \text{BESSIE})$ iff every sequence satisfies this sentence of $PL$.

If any sequence $\sigma$ is such that $\sigma(\text{VENUS})$ is above $\sigma(\text{BESSIE})$, then for each sequence $\sigma^*$, $\sigma^*(\text{VENUS})$ is above $\sigma^*(\text{BESSIE})$. So if a sentence of $PL$ has no inconstant variables, it is satisfied by each sequence or none. Thus, we can say that any such sentence is true iff it is satisfied by all sequences, and false iff it is satisfied by no sequence. This characterizes truth and falsity, for some sentences, in terms of the more basic and recursively characterized notion of satisfaction: $\sigma$ satisfies $\neg[COW(\text{BESSIE}) \& ABOVE(\text{VENUS}, \text{SADIE})]$ iff it is not the case that $\sigma$ satisfies both $COW(\text{BESSIE})$ and $ABOVE(\text{VENUS}, \text{SADIE})$. This condition is satisfied by all sequences or none. So the negated sentence is false or true, depending on whether or not Bessie is a cow and Venus is above Sadie. But this truth condition is characterized in terms of the satisfaction condition for the negated sentence, not in terms of the falsity/truth of $[COW(\text{BESSIE}) \& ABOVE(\text{VENUS}, \text{SADIE})]$, much less the truth/falsity of $COW(\text{BESSIE})$ and $ABOVE(\text{VENUS}, \text{SADIE})$.

In $PL$, the symbols ‘$\neg$’ and ‘&’ are not of types $<t, t>$ and $<t, <t, t>>$. That is why the sentences of $PL$ can include $[COW(x'') \& \neg ABOVE(x', x)]$ and $\neg[COW(x'') \& ABOVE(\text{VENUS}, x)]$. Sentential significance is specified in terms of sequences of d-entities, not in terms of truth. The variable-free sentences of $PL$ do not denote (truth) values of functions indicated by predicative constituents given arguments indicated by constants. Sentences of $PL$ specify conditions on ways of ordering d-entities. Indeed, each sentence is significant in and only in the following sense: it divides the ways of ordering a domain into those that meet a certain condition, and the rest.

In general, potential values of variables will not themselves determine whether they meet the condition imposed by $[GIVE(x, x', x'') \& \neg COW(x')]$. Given three d-entities, they might meet the condition or not, depending on which one is assigned to which variable. But a sequence of variable values is just the sort of thing that can meet this condition or not, depending on whether or not: the first d-entity of the sequence gave the second to the third, and the second is not a cow. One can say that each sentence of $PL$ is of type $<PLTF>$, for pre-linked thought-frame; see §4.2 of chapter one. One can add that ‘$\neg$’ and ‘&’ are of the types $<PLTF, PLTF>$ and $<PLTF, <PLTF, PLTF>>$. But whatever our notation, sequences are the only things that can satisfy a sentence of $PL$, regardless of how many or how few variables/constants the sentence includes.

---

13 For explicitness, order the constants of PL, and let each sequence $\sigma$ have the corresponding (finite) $n$-tuple of d-entities as a “zeroth” element; cp. Larson and Segal’s (1995) encoding Kaplan’s (1989) treatment of indexicals.

14 A true sentence can be conjoined with $COW(x)$ to form a sentence that is satisfied by some but not all sequences, and so neither true nor false. Even for purposes of specifying truth, satisfaction conditions are what matter.

15 One can say that $ABOVE(x', x'')$ is “tupled” by ordered pairs $(\alpha, \beta)$ such that for some sequence $\sigma$: $\sigma(x') = \alpha, \sigma(x'') = \beta$, and $\sigma$ satisfies $ABOVE(x', x'')$; $GIVE(x, x', x'')$ is “tupled” by ordered triples; etc. For any sentence $S$ of PL, some $n$ is such that only the first $n$ entities of a sequence $\sigma$ will be relevant to whether or not $\sigma$ satisfies $S$. But for any finite $n$, endlessly many sentences of $PL$ have more than $n$ variables. So sequences need to be unbounded, even given an upper bound on the adicities of atomic sentences.

14
2.3 Quantification via Sequences

Elementary quantification is easily introduced as follows: for any variable \( n \) and sentence \( S \) of PL, \( \exists n[S] \) is also a sentence of PL. Thus, \( \exists x[\text{ABOVE}(x', x)] \) and \( \exists x[\text{ABOVE}(x, x')] \) are sentences, as are \( \exists x'['\text{GIVE}(x, x', x')] \) and \( \exists x'[\text{COW}(x) \& \text{ABOVE}(x', x)] \). Let’s also permit elimination of doubled brackets, so that the last sentence can be respelled as \( \exists x'[\text{COW}(x) \& \text{ABOVE}(x', x)] \).

Like all sentences of PL, these new examples specify conditions on sequences. But for each sequence \( \sigma \), a sentence of the form \( \exists n[S] \) imposes a weak condition on \( \sigma \): some sequence that differs minimally from \( \sigma \), in the respect specified by \( n \), meets the condition imposed by \( S \). Tarski’s notion of a sequence variant, implicit in the plentitude of sequences, captures the relevant notion of minimal difference. For any sequence \( \sigma \) and variable \( n \), a sequence \( \sigma*\) is an \( n \)-variant of \( \sigma \)—more briefly, \( \sigma* \approx_n \sigma \)—iff \( \sigma* \) differs from \( \sigma \) at most with respect to the value of \( n \); where this means that either \( \sigma* \) does not differ at all from \( \sigma \) (i.e., \( \sigma* = \sigma \)), or \( \sigma* \) differs from \( \sigma \) only in the \( n \)th position. The relation indicated with ‘\( \sigma* \approx_n \sigma \)’ is reflexive, symmetric, and transitive. But note that \( \text{BVS} \sigma \) differs from \( \text{BVS} \sigma \) in two respects: \( \text{BVS} \sigma(X) \neq \text{BVS} \sigma(x) \), and \( \text{BVS} \sigma(x') \neq \text{BVS} \sigma(x) \). So \( \text{BVS} \sigma \) is neither an \( x \)-variant nor an \( x' \)-variant of \( \text{BVS} \sigma \). If \( \text{SVS} \sigma \) is just like \( \text{BVS} \sigma \) except that \( \text{SVS} \sigma(x) \) is Sadie the horse, then \( \text{SVS} \sigma \) is a mutant of \( \text{BVS} \sigma \); but these \( x \)-variants are not \( x' \)-variants.

Sequences alike at every position except the \( n \)th are \( n \)-variants. But if \( n \) and \( m \) are distinct variables, sequences that differ at the \( n \)th position cannot be \( m \)-variants.

Sequence \( \sigma \) satisfies a sentence of the form \( \exists n[S] \) iff some \( n \)-variant of \( \sigma \) satisfies \( S \). Put another way, \( \sigma \) satisfies \( \exists n[S] \) iff some sequence \( \sigma* \) such that \( \sigma* \approx_n \sigma \) is such that \( \sigma* \) satisfies \( S \). This treats ‘\( \exists n \)’ as a kind of quantification over sequences. For example, \( \sigma \) satisfies \( \exists x[\text{COW}(x)] \) iff some \( x \)-variant of \( \sigma \) satisfies \( \text{COW}(x) \). Hence, \( \text{SVS} \sigma \) satisfies \( \exists x[\text{COW}(x)] \), even though \( \text{SVS} \sigma(x) \) is Sadie the horse. For at least one sequence—\( \text{BVS} \sigma \) is an example—is such that it is an \( x \)-variant of \( \text{SVS} \sigma \), and its first element is cow. Obviously, \( \text{BVS} \sigma \) also satisfies \( \exists x[\text{COW}(x)] \); since \( \text{BVS} \sigma \) itself satisfies \( \text{COW}(x) \), and each sequence is an \( x \)-variant of itself. And so long as there is a cow in the domain, there will be many sequences whose first element is a cow; in which case, every sequence will satisfy \( \exists x[\text{COW}(x)] \), even if endlessly many sequences do not satisfy \( \text{COW}(x) \). But if there is no cow in the domain, no sequence will have a cow as its first element. In which case, no sequence will satisfy \( \text{COW}(x) \), and so no sequence will satisfy \( \exists x[\text{COW}(x)] \). Thus, \( \exists x[\text{COW}(x)] \) is satisfied by all sequences or none, depending on whether or not there is a cow in the domain.

Likewise, \( \sigma \) satisfies \( \exists x[\text{ABOVE}(x', x) \& \text{COW}(x)] \) iff some sequence \( \sigma* \) such that \( \sigma* \approx \sigma \) is such that: \( \sigma* \) satisfies \( \text{ABOVE}(x', x) \& \text{COW}(x) \); i.e., \( \sigma* \) satisfies \( \text{ABOVE}(x', x) \), and \( \sigma* \) satisfies \( \text{COW}(x) \); i.e., \( \sigma*(x') \) is above \( \sigma*(x) \), and \( \sigma*(x) \) is a cow. One can paraphrase this condition on sequences given two facts: if \( \sigma* \approx \sigma \), then \( \sigma*(x') = \sigma(x') \); and if \( \sigma*(x) \) is a cow, then being above \( \sigma*(x) \) is being above a cow. So \( \sigma \) satisfies \( \exists x[\text{ABOVE}(x', x) \& \text{COW}(x)] \) iff \( \sigma(x') \) is above a cow. Thus, despite the asterisks and squiggles in many specifications of this condition on sequences, the condition is simple: the second sequence element is above a cow.

Let’s abbreviate instances of ‘\( \_ \) satisfies \( S \) iff ...’ as instances of ‘\( \text{Sat} \{\_ \}, S \iff ... \)’, and abbreviate instances of ‘sequence \( \_ \) such that ...’ with instances of ‘\( \_ \) s ...’, as shown below.

\[
\text{Sat} \{\sigma, \exists x'[\exists x[\text{ABOVE}(x', x) \& \text{COW}(x)]]\} \iff \text{some } \sigma* : \sigma* \approx \sigma \text{ is such that } \\
\text{Sat} \{\sigma*, \exists x[\text{ABOVE}(x', x) \& \text{COW}(x)]\}
\]

Repeating the reasoning above, without the paraphrase, yields the following biconditional.

\[
\text{Sat} \{\sigma*, \exists x[\text{ABOVE}(x', x) \& \text{COW}(x)]\} \iff \text{some } \sigma** : \sigma** \approx \sigma* \text{ is such that } \\
\sigma**(x') \text{ is above } \sigma**(x), \text{and } \sigma**(x) \text{ is a cow}
\]
So the following sequence specification condition can be derived, given that ‘$\iff$’ is transitive.\(^\text{16}\)

\[
Sat\{\sigma, \exists x'[\exists x[\text{ABOVE}(x', x) \& \text{COW}(x)]]\} \iff \text{some } \sigma^*: \sigma^* \equiv_{v} \sigma \text{ is such that some } \sigma^{**}: \sigma^{**} \equiv_{v} \sigma \text{ is such that } \sigma^{**}(x') \text{ is above } \sigma^{**}(x) , \text{ and } \sigma^{**}(x) \text{ is a cow.}
\]

That is, $\sigma$ satisfies $\exists x'[\exists x[\text{ABOVE}(x', x) \& \text{COW}(x)]]$ iff some d-entity (that is the first element of some $x'$-variant of $\sigma$) is a above a cow. And this condition is satisfied by all sequences or none.

In chapter four, I return to the importance of \textit{deriving}—as opposed to merely formulating and paraphrasing—specifications of conditions on sequences. But to foreshadow, and to take a break from mere technicalia: while Tarski offered his semantics as an algorithm that specifies a certain relation that invented expressions bear to an idealized domain of entities, one can imagine a mind that already enjoys a \textit{richer} language of thought than PL. In particular, imagine a mind that can conduct derivations like those conducted in a Tarskian \textit{metalanguage}. Such a mind would not merely generate sentences of PL. It would be able to represent such sentences as having satisfaction conditions by using mental analogs of ‘\textit{Sat}\{\_, S\} \iff \ldots ’, sequence variables, ‘$\sigma^*: \sigma^* \equiv_{v} \sigma'$’, and other apparatus required to derive specifications of conditions on sequences.

One can hypothesize that competent speakers of a human language H likewise represent satisfaction conditions for sentences of H. Then while each true sentence would bear the same semantic relation to the domain—being satisfied by all sequences—true sentences could still be distinguished by virtue of being derivationally paired with distinct \textit{specifications of conditions} on sequences. In this sense, true sentences could differ in meaning, even if the domain does not include sentential meanings that true sentences have. Initially, it might seem crazy to suggest that ordinary speakers tacitly conduct Tarski-style derivations. But it may be worse to deny this hypothesis, yet still follow Davidson (1967a) in supposing that English has a Tarski-style semantics. One wants know else sentences of a human language H could be satisfied by sequences if competent speakers/signers of H \textit{don't} represent the relevant satisfaction conditions.

A lucky thinker might have concepts that are related to d-entities (and to each other) in the right way, so that the thinker enjoys a language of thought (e.g., PL) that has a Tarskian semantics that the thinker does not represent. But pronounceable expressions are presumably related to d-entities indirectly, via uses that are mediated by concepts; cp. Chomsky (195x). So even if some concepts have satisfaction conditions that go unrepresented, it is hard to see how words could have satisfaction conditions except via mentalistic specifications.\(^\text{17}\) In the end, I reject the hypothesis that ordinary speakers “internalize Tarski,” but not because I think children could not have the requisite mental language and use it as a meta-language. Animal minds are impressive; see, e.g., Gallistel (200x) on insect cognition. I have doubts about appealing to mental analogs of ‘Sat\{\_, S\}’.

But given the quantificational thoughts that ordinary humans enjoy, it is reasonable to posit mental analogs of quantification over sequence variants. So I take “psychologized” versions of Tarskian semantics (e.g., Larson and Segal [1995]) seriously; see chapter four.

At any rate, one can view \textit{derivations} of specifications of conditions on sequences—derivations of \textit{“Sat-sentences”}—as models of how a mind equipped with classical semantic

\(^{16}\text{See Larson and Segal (1995). Any system of derivation will require a version of ‘P iff Q, and Q iff R, so P iff R’, the sentential analog of ‘a = b and b = c, so a = c’. But while the Fregean biconditional ‘$\equiv$’ must be flanked by truth value denoters, the double arrow used here can be flanked by Tarskian sentences.}

\(^{17}\text{Davidson famously appealed to a notion of idealized interpretation as an alleged sufficient condition: a speaker’s word ‘cow’ is true of cows if her behavior would lead a certain kind of interpreter to adopt a theory that includes this hypothesis. But there is little reason for thinking that this condition is sufficient or ever met by actual humans.}
concepts might use an extant mental language to generate (i) hypotheses regarding the satisfaction conditions of foreign sentences, or (ii) stipulations regarding the satisfaction conditions of sentences generated by a new procedure implemented by that very mind. At least in case (ii), the new sentences could be articulable devices of communication, as opposed to instructions for how to build sentences of the older mental language. So again, I take this view seriously. And note that thinkers might agree on which sequences satisfy \textit{COW}(x), even if they think about cows differently. The new sentences—e.g., \( \exists x'[\exists x[\text{ABOVE}(x', x) \& \text{COW}(x)]] \)—would impose conditions on sequences indirectly, by virtue of being paired with mental analogs of ‘some \( \sigma^*: \sigma^*\models \cdot \sigma \) is such that some \( \sigma^{**}: \sigma^{**}\models \cdot \sigma \) is such that \( \sigma^{**}(x') \) is above \( \sigma^{**}(x) \), and \( \sigma^{**}(x) \) is a cow’. But depending on one’s goals, this might be perfectly acceptable; cp. Katz and Fodor (1963), Lewis (1972), Dummett (1976).

I’ll return to these issues. But let me get back to a few last technical details regarding PL. Given negation and the existential quantifier introduced above, we can define a universal quantifier. Or we can say that for any sequence \( \sigma \), variable \( n \), and sentence \( S \) of PL: \( \sigma \) satisfies \( \forall n S \) iff every \( n \)-variant of \( \sigma \) satisfies \( S \). That is, \( \sigma \) satisfies \( \forall n[S] \) iff every \( \sigma^*: \sigma^*\models^n \sigma \) is such that \( \sigma^* \) satisfies \( S \). For example, \( \text{Sat}\{\sigma, \forall x[\text{COW}(x)]\} \iff \forall \sigma^*: \sigma^*\models^n \sigma \) is such that \( \text{Sat}\{\sigma^*, \text{COW}(x)\} \). So \( \text{BSV} \sigma \) does not satisfy \( \forall x[\text{COW}(x)] \), even though \( \text{BSV} \sigma(x) \) is a cow. If even one d-entity is not a cow, then each sequence has an \( x \)-variant that does not satisfy \( \text{COW}(x) \), and no sequence satisfies \( \forall x[\text{COW}(x)] \). But if everything in the domain is a cow, each \( (x\text{-variant of}) \) each sequence satisfies \( \text{COW}(x) \). So \( \forall x[\text{COW}(x)] \) is satisfied by all sequences or none, depending on whether or not each d-entity is a cow. Likewise, \( \forall x[\text{COW}(x) \& \neg \text{BROWN}(x)] \)—and the equivalent \( \forall x[\text{COW}(x) \supset \text{BROWN}(x)] \)—is satisfied by all sequences or none, depending on whether or not each d-entity fails to be both a cow and not brown.

Sentences with both quantifiers, as in \( \forall x'[\exists x[\text{ABOVE}(x', x) \& \text{COW}(x)]] \), are already accommodated by the interpretive rules for instances of \( \forall n S \) and \( \exists n S \).

\( \text{Sat}\{\sigma, \forall x'[\exists x[\text{ABOVE}(x', x) \& \text{COW}(x)]]\} \iff \forall \sigma^*: \sigma^*\models^n \sigma \) is such that \( \text{Sat}\{\sigma^*, \exists x[\text{ABOVE}(x', x) \& \text{COW}(x)]\} \)

Since we just specified what it is for \( \sigma^* \) to satisfy \( \exists x[\text{ABOVE}(x', x) \& \text{COW}(x)] \), the following sequence specification condition can be derived.

\( \text{Sat}\{\sigma, x'[\exists x[\text{ABOVE}(x', x) \& \text{COW}(x)]]\} \iff \forall \sigma^*: \sigma^*\models^n \sigma \) is such that some \( \sigma^{**}: \sigma^{**}\models^n \sigma \) is such that \( \sigma^{**}(x') \) is above \( \sigma^{**}(x) \), and \( \sigma^{**}(x) \) is a cow.

That is, \( \sigma \) satisfies \( \forall x'[\exists x[\text{ABOVE}(x', x) \& \text{COW}(x)]] \) iff every \( x'\)-variant of \( \sigma \) is such that it has an \( x \)-variant whose second element is above its first, and its second element is a cow. Or more simply, a sequence satisfies this sentence of PL iff each d-entity is above a cow.

Instances of \( \forall n S \) impose a strong condition on sequences. If \( \sigma \) satisfies \( \forall n S \), then \( \sigma \) satisfies \( S \), and every other \( n \)-variant of \( \sigma \) satisfies \( S \). Instances of \( \exists n S \) impose a weak condition: \( \sigma \) can satisfy \( \exists n S \) without satisfying \( S \), so long as some \( n \)-variant of \( \sigma \) satisfies \( S \). Yet in either case, if \( S \) has no inconstant variables apart from \( n \), each of these conditions will be met by every sequence or none. So we can say that a sentence of PL is \textit{true} iff it is satisfied by all sequences. And since universal/existential quantification is construed in terms of strong/weak constraints on sequences, \( \forall x[\Phi(x)] \) implies \( \Phi(\text{BESSIE}) \), which implies \( \exists x[\Phi(x)] \); though not conversely.

---

18 At least if the sentence has no free (i.e., inconstant and unbound) variables. It is a matter of decision whether we count open sentences like \( \text{COW}(x) \) as true given a domain that includes only cows.
2.4 Semantic Recursion without Semantic Composition

The significance of PL sentences is characterized recursively in terms of satisfaction. Given how truth is related to satisfaction, each closed sentence of PL is true or false, depending on whether or not some (and hence every) sequence meets a certain condition. In this sense, each closed sentence $S$ of PL has a truth condition that can be specified with an instance of the following schema: $\text{True}(S) \iff F(\alpha)$. But in general, the sequences that satisfy a sentence are not composed of sequences that satisfy parts of the sentence; cp. Salmon (200x). One can say that sets of sequences are “$T$-meanings,” which do not compose. This is not a problem given Tarski’s aims. If one just needs a method that determines interpretations for generable formulae, as opposed to a theory of some natural generative procedure, one may not even need to specify the interpretations recursively in terms of constituents of the formulae; see Szabo (200x). But since expressions of PL are often viewed as models of human linguistic expressions, it is worth highlighting the distinction between being specifiable and being composite. For it may be that human languages are unlike PL in being less systemically connected to d-entities, but significant in way that is more demandingly compositional.

A sequence can satisfy $\text{cow}(x)$ without satisfying $\forall x[\text{cow}(x)]$. But the sequences that satisfy $\forall x[\text{cow}(x)]$—either all of them or none—are not composed of the sequences, if any, that satisfy $\text{cow}(x)$; and since $\forall x$ is not a sentence, it cannot be satisfied by any sequences. Perhaps even more obviously, the satisfiers of $\exists x[\text{cow}(x)]$ are not built up from whatever sequences satisfy $\text{cow}(x)$. A sequence can satisfy $\exists x[\text{cow}(x)]$ without satisfying $\text{cow}(x)$. We can say that the satisfiers of $\exists x[\text{cow}(x)]$ are determined by the satisfiers of $\text{cow}(x)$, given the rule for interpreting instances of $\exists nS$, and likewise $\forall x[\text{cow}(x)]$. This assumes that quantificational sentences of PL exhibit a kind of syntactic composition that corresponds to their recursive specification. In particular, one can view $\exists x[\text{cow}(x)]$ and $\forall x[\text{cow}(x)]$ as results of embedding the sentence $\text{cow}(x)$ in the frames $\forall x[\ldots]$ and $\forall x[\ldots]$, thereby forming larger sentences that have $\text{cow}(x)$ as a shared part. And specifications of conditions on sequences, mentalistic or orthographic, can be composite. Expressions of a metalanguage can exhibit part-whole relations that mirror syntactic composition in an object language. But sequences don’t compose.\(^{19}\)

Relatedly, embedding a Tarskian sentence in a sentential frame is not a notational variant of saturating a Fregean concept. The concept $\text{cow}(\_)$, which is not a sentence of PL, can be saturated by a concept of type $<e>$; and $\text{cow}(\_)$ can saturate the concept $\forall[\Phi(\_)]$, which is of type $<et, t>$. But $\forall x$ is not a Fregean concept of any type. It is a kind of prefix that extends one sentence into another. That is why $\forall x[\text{above}(x', x')]$ is a sentence with satisfiers—as opposed to a mere concatenation of two mismatched concepts, one of type $<et, t>$ and one of type $<e, et>$. Likewise, $\neg$ is not of type $<t, t>$ in PL. The satisfiers of $\neg\text{cow}(x)$ are determined by the satisfiers of $\text{cow}(x)$, given the interpretive rule for instances of $\neg S$. But the sequences that satisfy $\neg\text{cow}(x)$ are not composed of those that satisfy $\text{cow}(x)$; and $\neg$ cannot be satisfied by sequences. Attaching a prefix to $\text{cow}(x)$ differs from saturating $\text{cow}(\_)$ or the prefix. Prefixing turns one predicate of sequences into another. Saturating reduces the adicity of a concept.

Similarly, Tarski’s ampersand cannot be satisfied or saturated. From a Fregean perspective, this ampersand is not a concept. It is, at best, a syncategorematic device that can be used to form complex concepts. In this respect, $\&$ is like the variable extender ‘.’. The satisfiers of

\(^{19}\) Fregean Bedeutungen don’t compose either. The truth value of an ideal thought, of the form $\Phi(\alpha)$, is not built up from the function and entity indicated with $\Phi(\_)$ and $\alpha$. But Sinnen are supposed to exhibit part-whole relations, perhaps on analogy with interlocking pieces of a jigsaw puzzle, or valenced atoms that combine to form molecules.
Also think that real values cannot vary. But ignores the fact that in Fregean expressions used in thinking about how human languages are determined by the satisfiers of $\text{ABOVE}(\cdot')$ and $\text{COW}(\cdot)$, given the rule for interpreting instances of $[S & S^*]$; and a conjunctive sentence is built up from its parts. But sentences of PL do not have meanings that are determined by, much less composed of, the meanings of subsentential expressions like $\text{ABOVE}(_{-}, _{-})$ and the ampersand. One can say that nonsentential symbols of PL have a kind of derived significance, as potential sentence parts. But in PL, not even $\text{BESSIE}$ and $\text{COW}(_{-})$ are formatives that have significance of their own. The significance of $\text{COW}(\text{BESSIE})$, as a sentence of PL, is related to the contents of the Fregean concepts $\text{BESSIE}$ and $\text{COW}(_{-})$; and likewise for $[\text{ABOVE}(\cdot''), \text{BESSIE}] & \text{COW}(\cdot)]$. But to repeat, there is no Fregean counterpart of the Tarskian ampersand. And in PL, satisfaction by sequences is the only kind of significance that expressions (i.e., sentences) can have.

I have been describing PL as a possible generative system whose atomic concepts, which exhibit a saturated/unsaturated distinction, can be used to form mental sentences that contain bindable variables. To show that the resulting formulae still count as concepts—representations with which thinkers can think, as opposed to mere results of jamming variables into slots for genuine saturaters—one must show how the sentences of PL can have a systematic kind of significance. Tarski provided the resources for doing this. So we can see how, at least in principle, concepts like $\text{ABOVE}(_{-}, _{-})$ and $\text{COW}(_{-})$ could be used to form truth-allowable representations like $\forall x' \exists x [\text{ABOVE}(x', x) & \text{COW}(x)]$.

That was a major achievement. But crucially, $\forall x' \exists x [\text{ABOVE}(x', x) & \text{COW}(x)]$ exhibits a kind of composition that differs from the kind exhibited by $\text{COW}(\text{BESSIE})$. And one might suspect that with respect to composition, meanings are more like concepts. But even setting aside questions about the sense(s) in which meaning is compositional, one might want to consider variants of PL that allow for certain kinds of subsentential significance; cp. Lewis on grammars, as discussed in §2.3 of chapter zero. Human languages generate expressions that seem not to have analogs in PL—e.g., adverbial modifiers, indexicals, demonstratives, that-clauses, plural noun phrases, and quantificational sentences whose interpretations cannot be captured in terms of Tarskian quantification over sequences. This last topic was especially important in the development of formalisms for describing meanings. So let me move ahead with the quasi-historical review.

3. **Supplementing PL: Taking Tarski to Church**

In this section, I review Church’s (1941, 195x) extension of Tarski-style semantic theories. Church used his independently developed lambda calculus (see §… of chapter zero) to provide a semantics more like the one Frege envisioned for his Begriffsschrift, with many meaningful expressions having denotations or Bedeutungen. Church’s invented language, unlike Tarski’s, was also designed to permit representation of any computable function. So as Montague (1974) illustrated, Church’s toolkit provides more resources for modeling human languages in truth-theoretic terms. But unsurprisingly, this highlights questions concerning interpretations that human linguistics expressions do not exhibit.

3.1 **Truth Values, Entities, and Functions as Denotations**

In thinking about how human languages apparently differ from PL, one might focus on expressions used in reference and quantification. A name like ‘Bessie’ seems somehow like a Fregean concept of type $<e>$, with which one can think about a particular thing, at least if one ignores the fact that individuals share names. Names can be modeled with variables whose values cannot vary. But this suggests a special kind of significance—viz., denoting. One might also think that real variables are for binding, and that phrases like ‘every cow’ are binders of the type $<et, t>$. That raises questions, to which I return in chapter five, about ‘every Bessie I know’;
cp. Burge (197x). But given a categorematic conception of names and quantifiers, predicates presumably exhibit a kind of significance that sentences of PL do not. If ‘is brown’ can combine with expressions of type <e> or <et, t> form a sentence, then ‘is brown’ is like BROWN(_) and COW(_) in being of type <et>. Still, even if one posits unsaturated concepts, one might want to avoid Frege’s appeal to undenotable unsaturateds. Church (1941, 195x) showed how to extend a Tarskian language, in a way that reflects these considerations, by introducing combinable denoters of various types.

I will not, however, follow Church’s presentation in detail. My aim is to show how idealized minds might introduce a language of thought that has a Church-style semantics, given PL and a few additional concepts. For one might think children have or acquire a mental language that lets them specify denotations (a.k.a. “semantic values”) for all meaningful expressions, as opposed to merely specifying satisfaction conditions for sentences (open or closed). Indeed, this is tempting construal of much recent work in semantics. I also want to describe this extended language—call it CPL—in a way that highlights contrasts with PL, while noting that CPL is a Fregean language that avoids the difficulties discussed in section two. This will serve as a model for how human minds might come to have a less powerful language of thought (SMPL), described in chapters five and six. Specialists can keep skimming. But it is easy to forget how Church-style denotation of Fregean functions (in extension) is related to Tarski’s notion of satisfaction. And later, it will be important to have been clear about this.

Like PL, CPL is an idealized generative procedure. Though for purposes of interpreting the generable expressions, let the domain of PL be a subset of the domain of CPL. The domain of PL is the domain of entities, of type <e>, for CPL. Let CPL also have a domain of truth values, of type <t>, exhausted by two entities T and ⊥. It does not matter if truth values are distinctive abstracta, or two entities of type <e> chosen at random. What matters is that there are exactly two such values, so that any function of type <e, t> maps each entity of type <e> to T or ⊥. In which case, any such function maps an entity to ⊥ if and only if that function does not map that entity to T. Likewise, each function of type <t, t> maps each truth value to T or ⊥; each function of type <et, t> maps each function of type <e, t> to T or ⊥; etc. And let CPL have a domain of functions, comprising (sub)domains of endlessly many types, recursively characterized in terms of the two basic domains of type <e> and <t>; given a domain of type <a> and a domain of type <β>, CPL also has a domain of functions of type <α, β>.

The domain for CPL thus includes subdomains that mirror the Fregean hierarchy of conceptual types; see §1.5 above. Things of type <<α> differ from concepts of type <α>, apart from concepts of type <e>, which are also entities. But I’ll adopt the standard if potentially misleading practice of using the same typological notation for subdomains and concepts. However, the domain of CPL does not include any “unsaturateds” in Frege’s intended sense (see §1.1 above). The functions that are stipulated to be in the domain for CPL—given the entities and truth values—are denotable sets of ordered pairs.²⁰ So each thing in the domain of CPL is denotable, and one can extend PL to include corresponding denoters.

Let the sentences of CPL include all the sentences of PL. And let’s say that for any sentence S of PL, T-value {S} is a denoting expression of CPL. As this suggests, the relevant notion of denotation will be sequence relative, since denoting expressions may contain unbound variables. Given a closed sentence like ∃X[COW(X)], one might simply say that the corresponding

---

²⁰ These sets (E-functions) may be specified in terms of procedures (I-functions) that we can think about. Though one also can interpret CPL in terms of a hierarchical domain of procedures; see §2.1 of chapter zero.
truth value is \( T \) or \( \perp \): \text{T-VALUE}\{\exists x [\text{cow}(x)]\} denotes \( T \) if \( \exists x [\text{cow}(x)] \) is satisfied by each sequence; otherwise \text{T-VALUE}\{\exists x [\text{cow}(x)]\} denotes \( \perp \). But we don’t want to say that \( \text{cow}(x) \) has a truth value, once and for all, much less the same truth value as \( \exists x [\text{cow}(x)] \) or \( \forall x [\text{cow}(x)] \). Still, we want CPL to generate the expression \text{T-VALUE}\{\text{cow}(x)\}, so that \( \text{cow}(x) \) can be part of a representation of a corresponding function of type \( <e, t> \). So let’s say that denotation is sequence relative, allowing for expressions whose denotation is constant across sequences.

We can think of sequences as ways of assigning values to unbound variables, and say that relative to any sequence \( \sigma \): the denotation of \text{T-VALUE}\{S\} is \( T \) if \( \sigma \) satisfies \( S \); otherwise, the denotation of \text{T-VALUE}\{S\} is \( \perp \). Since a closed sentence is satisfied by all sequences or none, the denotation of \text{T-VALUE}\{\exists x [\text{cow}(x)]\} is the same relative to each sequence: \( T \) or \( \perp \), depending on whether or not the domain of entities includes a cow. But relative to each sequence \( \sigma \), the denotation of \text{T-VALUE}\{\text{cow}(x)\} is \( T \) if \( \sigma \) satisfies \( \text{cow}(x) \), and otherwise \( \perp \). Put another way, for any assignment of a value to the variable in \text{T-VALUE}\{\text{cow}(x)\}, this expression denotes \( T \) or \( \perp \), depending on whether or not the assigned value is a cow. So if some but not all entities of type \( <e> \) are cows, \text{T-VALUE}\{\text{cow}(x)\} may denote \( T \) relative to \( \sigma \) and denote \( \perp \) relative to some other sequence \( \sigma^* \). Likewise, the denotation of \text{T-VALUE}\{\text{above}(x’, x)\} relative to \( \sigma \) will depend on whether or not \( \sigma \) satisfies \text{above}(x’, x): \( T \) if \( \sigma(x’) \) is above \( \sigma(x) \), and otherwise \( \perp \).

Each constant of PL is a denoter of CPL. But let’s continue to say that sequences map constants/denoters to entities. Relative to any sequence \( \sigma \), \text{Bessie} denotes \( \sigma(\text{Bessie}) \)—a.k.a. Bessie the cow. And for any sequences \( \sigma \) and \( \sigma^* \), \( \sigma(\text{Bessie}) = \sigma^*(\text{Bessie}) \).

In section three, I used notation like the following to specify conditions on sequences.

\[
\text{Sat} \{\sigma, \exists x [\text{cow}(x)]\} \iff \text{some } \sigma^* : \sigma^* = \sigma \text{ such that } \text{Sat} \{\sigma^*, \text{cow}(x)\}
\]

\[
\text{Sat} \{\sigma^*, \text{cow}(x)\} \iff \sigma^*(x) \text{ is a cow}
\]

\[
\text{Sat} \{\sigma, \exists x [\text{cow}(x)]\} \iff \text{some } \sigma^* : \sigma^* = \sigma \text{ such that } \sigma^*(x) \text{ is a cow}
\]

Recall that \( \text{Sat} \{\sigma, S\} \iff \ldots \) abbreviates \( \sigma \text{ satisfies } S \text{ iff } \ldots \), and \( \text{some } \sigma^* : \sigma^* = \sigma \) is short for \( \text{some sequence } \sigma^* \text{ such that } \sigma^* = \sigma \). Now let \( \text{Den} \{D, \sigma\} = \ldots \) abbreviate ‘the denotation of \( D \) relative to \( \sigma \) is \( \ldots \)’; and let \( \text{T/\perp}\ldots \) be short for ‘\( T \text{ if } \ldots \) and \( \perp \text{ otherwise} \’, as shown below.

\[
\text{Den} \{\text{Bessie}, \sigma\} = \sigma(\text{Bessie}) = \text{Bessie}
\]

\[
\text{Den} \{\text{T-VALUE}\{\text{cow}(x)\}, \sigma\} = \text{T/\perp}[\text{Sat} \{\sigma, \text{cow}(x)\}] = \text{T/\perp}[\sigma(x) \text{ is a cow}]
\]

Read ‘\( \text{T/\perp}[\sigma(x) \text{ is a cow}] \)’ as follows: \( T \) or \( \perp \), depending on whether not \( \sigma(x) \) is a cow.

The sentences of PL correspond to sequence-relative denoters of CPL, and \text{T-VALUE}\{\ldots\} can be viewed as an expression frame. But there are no instances of \text{T-VALUE}\{\text{T-VALUE}\{\ldots\}\}; and truth values do not have truth values. We can, however, introduce new denoting expressions as follows: for any variable \( \nu \) of type \( <\alpha> \) and denoter \( D \) of type \( <\beta> \), the expression \( \lambda \nu. D \) is a denoting expression of type \( <\alpha, \beta> \). Since \text{T-VALUE}\{\text{cow}(x)\} and \text{T-VALUE}\{\text{above}(x’, x)\} are denoting expressions of CPL, so are \( \lambda x. \text{T-VALUE}\{\text{cow}(x)\} \) and \( \lambda x. \text{T-VALUE}\{\text{above}(x’, x)\} \) and \( \lambda x. \lambda x. \text{T-VALUE}\{\text{above}(x’, x)\} \). The idea is that these lambda expressions denote functions that can be specified recursively as indicated below.

\[
\text{Den} \{\lambda x. \text{T-VALUE}\{\text{cow}(x)\}, \sigma\} = \text{the smallest function } F \text{ of type } <e, t> \text{ such that for each } \sigma^* : \sigma^* = \sigma, F \text{ maps } \sigma^*(x) \text{ to } \text{Den} \{\text{T-VALUE}\{\text{cow}(x)\}, \sigma^*\}
\]

Relative to any sequence \( \sigma \), \( \lambda x. \text{T-VALUE}\{\text{cow}(x)\} \) denotes the smallest function of type \( <e, t> \) that maps each entity in the relevant domain (i.e., each d-entity) to the truth value assigned to \text{cow}(x) given that entity as the value of the variable. For each sequence \( \sigma \) and d-entity \( \delta \), \( \sigma \)
has exactly one $x$-variant that assigns $\delta$ to $x$. So given that $\text{Den}\{\text{T-VALUE}\{\text{COW}(x)\}, \sigma^*\} = T/\perp[\sigma^*(x) \text{ is a cow}]$, the denotation of $\lambda x.\text{T-VALUE}\{\text{COW}(x)\}$ can be described as follows.

$$\text{Den}\{\lambda x.\text{T-VALUE}\{\text{COW}(x)\}, \sigma\} = \text{the smallest function } F \text{ of type } <e, t> \text{ such that for each } \delta, F \text{ maps } \delta \text{ to } T/\perp[\delta \text{ is a cow}]$$

This function—the cow-function, for short—maps Bessie to $T$ and Venus to $\perp$. So for any sequence $\sigma$, the cow-function maps $\text{Den}\{\text{BESSIE}, \sigma\}$ to $T$ and $\text{Den}\{\text{VENUS}, \sigma\}$ to $\perp$.

With more than one variable, the pattern is repeated.

$$\text{Den}\{\lambda x'\lambda x.\text{T-VALUE}\{\text{ABOVE}(x', x)\}, \sigma\} = \text{the smallest function } F' \text{ of type } <e, \mathbf{et}> \text{ such that for each } \sigma^* : \sigma^* =_x \sigma,$$

$$F' \text{ maps } \sigma^*(x') \text{ to } \text{Den}\{\text{T-VALUE}\{\text{ABOVE}(x', x)\}, \sigma^*\}$$

The (constant) denotation of $\lambda x'\lambda x.\text{T-VALUE}\{\text{ABOVE}(x', x)\}$ is thus specified in terms of the (inconstant) denotation of $\lambda x.\text{T-VALUE}\{\text{ABOVE}(x', x)\}$, which is specified in terms of the (inconstant) denotation of $\text{T-VALUE}\{\text{ABOVE}(x', x)\}$.

$$\text{Den}\{\lambda x.\text{T-VALUE}\{\text{ABOVE}(x', x)\}, \sigma^*\} = \text{the smallest function } F \text{ of type } <e, t> \text{ such that for each } \sigma^*:\sigma^* = x \sigma^*,$$

$$F \text{ maps } \sigma^*(x') \text{ to } \text{Den}\{\text{T-VALUE}\{\text{ABOVE}(x', x)\}, \sigma^*\}$$

So given a specification of the (inconstant) denotation of $\text{T-VALUE}\{\text{ABOVE}(x', x)\}$, the denotation of $\lambda x.\text{T-VALUE}\{\text{ABOVE}(x', x)\}$ can be characterized as shown below.

the denotation of $\lambda x.\text{T-VALUE}\{\text{ABOVE}(x', x)\}$ can be characterized as shown below.

$$\text{Den}\{\lambda x.\text{T-VALUE}\{\text{ABOVE}(x', x)\}, \sigma^*\} = \text{the smallest function } F \text{ of type } <e, t> \text{ such that for each } \sigma^*:\sigma^* = x \sigma^*,$$

$$F \text{ maps } \sigma^*(x') \text{ to } \text{Den}\{\text{T-VALUE}\{\text{ABOVE}(x', x)\}, \sigma^*\}$$

Or more briefly, and exploiting the fact that $\sigma^*(x') = \sigma^*(x')$ if $\sigma^*$ and $\sigma^*$ are $x$-variants,

$$\text{Den}\{\lambda x.\text{T-VALUE}\{\text{ABOVE}(x', x)\}, \sigma^*\} = \text{the smallest function } F \text{ of type } <e, t> \text{ such that for each } \delta, F \text{ maps } \delta \text{ to } T/\perp[\sigma^*(x') \text{ is above } \delta].$$

In which case, the denotation of $\lambda x'\lambda x.\text{T-VALUE}\{\text{ABOVE}(x', x)\}$ can be characterized as the smallest function $F'$ of type $<e, \mathbf{et}>$ such that

for each $\sigma^*:\sigma^* = x \sigma^*$, $F'$ maps $\sigma^*(x')$ to the smallest function $F$ of type $<e, t>$ such that for each $\delta$, $F$ maps $\delta$ to $T/\perp[\sigma^*(x') \text{ is above } \delta]$,

or more simply,

for each $\delta$, $F$ maps $\delta$ to the smallest function $F$ of type $<e, t>$ such that for each $\delta$, $F$ maps $\delta$ to $T/\perp[\delta' \text{ is above } \delta]$.

This function $F'$ maps Bessie to a function $F$ that maps each $d$-entity $\delta$ to $T$ or $\perp$ depending on whether or not Bessie is above $\delta$. So for any $\sigma$, $\text{Den}\{\lambda x'.\lambda x.\text{T-VALUE}\{\text{ABOVE}(x', x)\}, \sigma\}$ maps $\text{Den}\{\text{BESSIE}, \sigma\}$ to a function that maps $\text{Den}\{\text{VENUS}, \sigma\}$ to $T$ if Bessie is above Venus, and otherwise maps $\text{Den}\{\text{VENUS}, \sigma\}$ to $\perp$. By contrast,

$$\text{Den}\{\lambda x.\text{T-VALUE}\{\text{ABOVE}(x', x)\}, \sigma\} = \text{the smallest function } F \text{ of type } <e, \mathbf{et}> \text{ such that for each } \delta, F \text{ maps } \delta \text{ to the function } F' \text{ of type } <e, t> \text{ such that for each } \delta, F' \text{ maps } \delta' \text{ to } T/\perp[\delta' \text{ is above } \delta].$$

This function maps Bessie to a function that maps each entity $\delta'$ to $T$ or $\perp$ depending on whether or not $\delta'$ is above Bessie. So for any $\sigma$, $\text{Den}\{\lambda x.\lambda x'.\text{T-VALUE}\{\text{ABOVE}(x', x)\}, \sigma\}$ maps $\text{Den}\{\text{BESSIE}, \sigma\}$ to a function that maps $\text{Den}\{\text{VENUS}, \sigma\}$ to $T$ if Venus is above Bessie, and otherwise maps $\text{Den}\{\text{VENUS}, \sigma\}$ to $\perp$. 

22
3.2 Lambdas and Quantification over Sequences

The details just reviewed, somewhat tediously, illustrate the parallels between Church’s lambda-binders and Tarski’s quantifiers. And I want to highlight this point, which will be quite important in chapter six, before turning to combinations of CPL denoters.

If one viewed λ as a primitive symbol, unrelated to the core Tarskian idea of quantifying over sequence variants, it would be hard to see how—except by stipulation or magic—prefixing T-VALUE {\textit{cow}(x)} with ‘λx.’ could create a function denoter. It would be even harder to see how prefixing T-VALUE {\textit{above}(x’, x)} with two lambda binders could, depending on the order, create either of two function denoters. One wants to know how embedding \textit{cow}(x) in λx.T-VALUE{…} could have such an effect in any mind, ideal or natural—and likewise for embedding \textit{above}(x’, x) in λx’.λx.T-VALUE{…}, give(x”, x’, x) in λx”’λx’.λx.T-VALUE{…}, etc. Any sense of mystery evaporates, though, once one sees the relation between lambda abstraction and quantification over sequence variants (e.g., ‘some σ*:σ*=σ’).

We don’t know how natural minds implement such quantification, or how they represent functions (or entities). But given PL as a language of thought that can be extended in modest ways, forming the complex denoting expressions of CPL would not be a huge leap forward, despite initial appearances. This raises the question of whether PL already creates possibilities that are not open to natural human thought, even if such thought goes beyond PL in other ways. But lambda abstraction can be viewed as a way of unleashing the power of quantifying over sequence variants within a language that extends PL only in a few modest ways. And for Church (194x), the neo-Fregean application of his lambda calculus—recursive specification of functions that mirror Frege’s hierarchy of ideal concepts, rooted in the basic types <e> and <t>—merited only a passing mention (p. xx) as a special but not especially interesting case of how a certain class of functions can be specified with lambda expressions, given some simple primitives. Church invented his calculus to offer proofs concerning the entire space of Turing-computable functions; see §2.1 of chapter zero. Correlatively, lambdas unleash significant power in modest extensions of PL.

The flip side of this point is that lambda abstraction is not an alternative to quantification over sequence variants. It is a way of putting such quantification to work. This is no complaint. In chapter six, I posit a mental language that employs a restricted version of ‘some σ*:σ*=σ’ to generate certain monadic concepts from monadic and dyadic atoms. If appeal to quantification over sequence variants is unavoidable, in some form, we may as well make the most of whatever form we have to posit—even if we can’t yet say how natural minds implement some version of the trick whose explicit formulation required Tarski’s mind. So I have no objection to positing a language of thought with restricted analogs of lambdas. But what further mental resources do we need posit to account for word meanings, how they compose, and the cognitive productivity of expressible concepts? In particular, do meanings exhibit more than a smidgeon of dyadicity? If not, do we need to posit words that fetch concepts of \textit{functions} that map entities to truth values? If not, what do we need to assume instead, and what are the implications for how meaning is related to truth?

Such questions are not easily answered. But suppose that given a certain mentalese, the sparsest descriptively adequate form of quantification over sequence variants would yield representational capacities that humans do not naturally have. Then prima facie, human minds do not employ that mentalese. In thinking about the implications of this point, for theories of human languages, it is worth remembering two others. First, the power that ‘some σ*:σ*=σ’ unleashes depends partly on the number of variables that can be left open in a sentence. Second, the history
of semantics makes it easy to assume (even if one shouldn’t) that word meanings can be polyadic in ways that create multiple openings for abstraction. Indeed, Frege’s invention of modern logic—and his appeal to truth values as Bedeutungen of ideal thoughts (Gedanken)—was interwoven with his insight, which Church preserved, that apparent monadicity can be reconstructed as a special case of polyadicity. And it is worth pausing to be clear about this.

Instead of adopting a subject-predicate conception of thought—in which monadic concepts play a central role, but PRECEDES-EVERY-NUMBER-TAT-HAS-A-PREDECESSOR presents serious difficulties—Frege urged a function-argument conception of thought that initially seems to leave no place for traditional monadic concepts. Note that the predecessor function is usually specified with one variable: \( P(x) = x - 1; \lambda x.x - 1; \) etc. The dyadic character of the underlying concept, \text{PREDECESSOR}(\_, \_) is reflected by the fact that functions map inputs onto outputs; cp. \text{FATHER-OF}(\_, \_). The addition function is usually specified with two variables, while the underlying concept is intuitively triadic: \text{SUM}(\_, \_, \_). This raises the question of how to represent (the contents of) monadic concepts as functions. In the case of \text{NUMBER}(\_), Frege suggested a higher-order dyadic reanalysis; see §1.4 above. Perhaps \text{COW}(\_), \text{CALF}(\_), and \text{FARMER}(\_) are also better represented in terms of dyadic relations that individuals bear to times or events. But some concepts, including \text{NUMBER}(\_) itself, classify entities. So what do the corresponding functions map entities onto?

Frege’s answer is that ideal concepts indicate relations to truth values. The ideal concept \text{PREDECESSOR}(\_, \_) applies to ordered pairs of the form \( \langle (x, x'), t \rangle \); where the second element is a truth value, the first is an ordered pair of numbers, and the second element is \( T \text{ iff } x \) is the predecessor of \( x' \). Likewise, \text{SUM}(\_, \_) applies to ordered pairs \( \langle (x, x', x''), t \rangle \) such that \( t = T \text{ iff } x'' \) is the sum of \( x \) and \( x' \). This makes room for an analog of \text{COW}(\_), since \text{COW}(\_, \_) can apply to pairs \( \langle x, t \rangle \) such that \( t = T \text{ iff } x \) is a cow. This idea is reflected in the semantics for CPL, without problematic appeal to unsaturateds: \text{COW}(x) is used to form \( \lambda x.\text{T-VALUE}\{\text{COW}(x)\} \), whose denotation is a function whose extension is a set of entity/truth-value pairs.

One might speculate that humans naturally employ, at least for purposes of linguistic expression and comprehension, a system of thought that is fundamentally polyadic in this sense. As we’ll see in chapter five, many facts tell against this hypothesis. Still, one can at least imagine thinkers who naturally introduce concepts that denote functions that map entities to truth values. Given unsaturated concepts but no variables, a thinker might be able to form the thought \( \text{EVERY}[\text{ABOVE}(\_, \text{BESSIE}), \text{COW}(\_)] \), yet be unable to form concepts like \( \text{ABOVE}(\text{VENUS}, \_) \)—much less \( \text{COW}-\text{THAT}-\text{VENUS-IS-ABOVE}(\_), \text{COW}-\text{THAT}-\text{CHASED-EVERY-DOG}(\_), \text{SAW}-\text{A}-\text{COW}-\text{THAT}-\text{CHASED-EVERY-DOG}(\_), \text{etc.} \) But given CPL, a mind could use concepts like \( \text{ABOVE}(\_, \_) \) and \( \text{COW}(\_) \) to introduce complex denotes like \( \lambda x.\lambda x'.\text{T-VALUE}\{\text{ABOVE}(x', x)\} \), \( \lambda x'.\lambda x.\text{T-VALUE}\{\text{ABOVE}(x', x)\} \), \( \lambda x'.\lambda x.\text{T-VALUE}\{\text{ABOVE}(x', x) & \text{COW}(x)\} \), etc. This could be useful, since denotes like \( \text{BESSIE} \) and \( \lambda x.\text{T-VALUE}\{\text{COW}(x)\} \) can be combined in CPL.

### 3.3 Frege in Church: Polyadicity and Applying Functions

A mere pair of denoting concepts is not a concept, much less a thought. But in CPL, denoters of types \( <e, t> \) and \( <t> \) can combine to form denoters of type \( <t> \). More generally, if \( D \) is a denoting expression of type \( <\alpha, \beta> \) and \( D^* \) is a denoting expression of type \( <\alpha> \), then \( D \) can combine with \( D^* \) to form a complex denoting expression of the type \( <\beta> \). And relative to any sequence \( \sigma \), the denotation of the complex is the result of applying the function that is the denotation of \( D \) to the denotation of \( D^* \). Thus, “Apply Function”

\[
\text{(AF)} \quad \text{Den} \{ [D_{\alpha, \beta}, D^*_{<\alpha>}], \sigma \} = \text{Den} \{ D_{\alpha, \beta}, \sigma \} (\text{Den} \{ D^*_{<\alpha>}, \sigma \})
\]
is an interpretation principle for CPL. But note that (AF) differs from saturation in Frege’s sense.

In this metalanguage, $\Phi(\alpha)$ is the value of the function $\Phi$ given the argument $\alpha$. So if we take functions to be sets/procedures—abstract things, as opposed to Fregean unsaturated—then $\Phi(\alpha)$ is importantly different than the Fregean thought $\text{cow}(\text{Bessie})$, in which Bessie saturates $\text{cow}(\_)$ . Denoters cannot combine via saturation. Since $\lambda x.\text{T-VALUE}\{\text{cow}(x)\}(\text{Bessie})$ and $\text{T-VALUE}\{\text{cow}(\text{Bessie})\}$ have the same denotation, the former might be replaced with the latter in a derivation, suggesting an analogy to saturation. But $\lambda x.\text{T-VALUE}\{\text{cow}(x)\}$ is still a denoter, not a general concept like $\text{cow}(\_)$ . So Bessie cannot saturate $\lambda x.\text{T-VALUE}\{\text{cow}(x)\}$. This is no objection to CPL. Languages like PL and CPL are designed to overcome the limitations, described in section one, of saturation without variables. So it is no surprise if combination in CPL is not interpreted as saturation.

Indeed, CPL permits construction of $\lambda x'.\lambda x.\text{T-VALUE}\{\text{above}(x', x)\}(\text{Venus})$. This expression denotes the $<e, t>$ function that maps an entity to $T$ iff Venus is above that entity. Thus, CPL permits construction of an analog of $\lambda x'.\lambda x.\text{T-VALUE}\{\text{above}(\text{Venus}, x)\}(\text{Venus})$ as the co-denoting $\lambda x.\text{T-VALUE}\{\text{above}(\text{Venus}, x)\}$. But the latter expression is not formed by saturating $\lambda x'.\lambda x.\text{T-VALUE}\{\text{above}(x', x)\}$ with $\text{Venus}$. On the contrary, $\lambda x'.\lambda x.\text{T-VALUE}\{\text{above}(x', x)\}(\text{Venus})$ cannot be saturated, and that is why it merited introduction: the lambda expression denotes a function that maps Venus to the intended Bedeutung of $\lambda x'.\lambda x.\text{T-VALUE}\{\text{above}(x', x)\}(\text{Venus})$. Instead of simply assuming that Venus can skip over the first/rightmost slot of a polyadic concept and saturate another—or stipulating that Bessie can be extracted from $\lambda x'.\lambda x.\text{T-VALUE}\{\text{above}(\text{Venus}, x)\}(\text{Venus})$, leaving a saturatable concept—CPL provides a possible mechanism for generating a simulacrum of $\lambda x'.\lambda x.\text{T-VALUE}\{\text{above}(\text{Venus}, x)\}$ is formed by building $\lambda x'.\lambda x.\text{T-VALUE}\{\text{above}(x', x)\}(\text{Venus})$ and then rewriting this denoter in accord with an obvious rule for how to re-represent the value of a function given an argument.

If saturation is the only “direct” mode of combining concepts, then combining denoters and interpreting the results as Church-combinations is as “indirect” as combining monadic concepts and interpreting the results as M-junctions; see note 20 of chapter one). Not that there’s anything wrong with that. Recall that one might analyze the M-junction $\text{brown}^\sim\text{cow}(\_)$ as the result of twice saturating a concept of conjunction—$\cap\{\Phi(\_), (\_)\}$—of type $<\text{et}, <\text{et}, \text{et}>>$. Correlatively, the instruction $\text{M-join}[\text{fetch}@\text{cow'}, \text{fetch}@\text{brown'}]$ might be spelled out in terms of two instructions to saturate, at least if such instructions are generable.

Saturate[Saturate[fetch@\&, fetch@\text{cow'}, fetch@\text{brown'}]]

Executing the embedded step this mini-program would be a way of forming $\cap[\_, \text{cow}(\_)\]$. And this concept of type $<\text{et}, \text{et}>$ could be saturated with $\text{brown}(\_)$.

Likewise, one might analyze $\text{T-VALUE}\{\text{cow}(\text{Bessie})\}$ as the result of twice saturating a dyadic concept of application: $\text{APPLY}(\lambda x.\text{T-VALUE}\{\text{cow}(x)\}, \text{Bessie})$; where $\text{APPLY}(\_, \_)$ exhibits the abstract type $<D<\alpha>, <D<\alpha, \beta>, D<\beta>>$, and so saturating $\text{APPLY}(\_, \_)$ with two suitable denoters creates a third. The instruction Saturate[fetch@\text{cow'}, fetch@\text{Bessie'}] will not be executable if $\lambda x.\text{T-VALUE}\{\text{cow}(x)\}$ is the only concept at the ‘cow’ address. But a mind equipped with CPL might generate instructions that call for application rather than saturation: Apply[fetch@\text{cow'}, fetch@\text{Bessie'}]. And the instruction might be spelled out in terms of two instructions to saturate, given the concept Apply[fetch@\text{Bessie'}].
Executing the embedded complex instruction would be a way of forming $\text{APPLY}(\_, \text{BESSIE})$, a concept of the abstract type $<\text{D}_e, \beta, \text{D}_p>$. If the lexical address for ‘cow’ is shared by $\text{COW}(\_)$ and $\lambda x. \text{T}-\text{VALUE}\{\text{COW}(x)\}$, there are two ways of executing the lexical instruction $\text{fetch}@\text{‘cow’}$; see §4.2 of chapter one. But since $\text{APPLY}(\_, \_)$ can only be saturated by denoters, executing the program above would require that $\lambda x. \text{T}-\text{VALUE}\{\text{COW}(x)\}$ be the concept fetched via ‘cow’.

A rewriting instruction might then replace $\text{APPLY}(\lambda x. \text{T}-\text{VALUE}\{\text{COW}(x)\}, \text{BESSIE})$ with $\text{T}-\text{VALUE}\{\text{COW}(\text{BESSIE})\}$. For simple examples, one need not appeal to $\text{APPLY}(\_, \_)$. If $\text{COW}(\_)$ and $\lambda x. \text{T}-\text{VALUE}\{\text{COW}(x)\}$ reside at the same lexical address, then executing the instruction $\text{Saturate}[\text{fetch}@\text{‘cow’}, \text{fetch}@\text{‘Bessie’}]$ is still a way of forming $\text{COW}(\text{BESSIE})$. But the denoter $\lambda x. \text{T}-\text{VALUE}\{\text{ABOVE}(\text{BESSIE}, x)\}$ can be formed by saturating $\text{APPLY}(\_, \text{BESSIE})$ with the denoter $\lambda x'. \lambda x. \text{T}-\text{VALUE}\{\text{ABOVE}(x', x)\}$ and rewriting the result. Still, $\text{BESSIE}$ cannot saturate $\lambda x'. \lambda x. \text{T}-\text{VALUE}\{\text{ABOVE}(x', x)\}$ even if it can saturate the first position of $\text{APPLY}(\_, \_)$.

I don’t think that humans understand linguistic expressions by employing the concept $\text{APPLY}(\_, \_)$ along with Church-style denoters of functions/sets and capacities for rewriting such denoters. But just as one might hypothesize that children use a mental analog of PL to specify satisfaction conditions for human language sentences, one can hypothesize a mental analog of CPL. Moreover, my own proposal also involves constructing simulacra of Fregean concepts, by means of operations other than saturation, and letting lexical addresses be shared by concepts of different types. So it is important to be clear that in offering a Church-style reconstruction of Frege’s semantics, one appeals to operations other than saturation. The question is which further operations are needed, not whether further operations are needed.

I think that human languages invoke conjunction operations, that they do not invoke function-application, and that such languages are massively monadic as opposed to massively denotational. But this is an empirical hypothesis. One might have thought that since combining denoters cannot yield truth-evaluable thoughts, words like ‘cow’ cannot fetch denoting concepts that combine with others. This is not my view. I don’t think that ‘cow’ (or any word) fetches a denoter. But if thoughts can denote truth values, then denoters of the right sort can be truth-evaluable. Perhaps each thought needs at least one nondenoting constituent, as Frege suggested. But CPL can systematically generate denoters that can saturate $\text{APPLY}(\_, \_)$—and thereby form analogs of Fregean thoughts whose constituents include concepts like $\text{COW}(\_)$, as opposed to $\text{APPLY}(\_, \_)$ and a denoter like $\lambda x. \text{T}-\text{VALUE}\{\text{COW}(x)\}$.

Indeed, my hypothesis that linguistic meanings direct construction of predicates can be viewed as a variant of the idea that linguistic meanings direct construction of CPL denoters. In principle, a semanticist can reject the latter hypothesis and still say that the “semantic value” of the word ‘cow’ is the function $\lambda x. \text{T}-\text{VALUE}\{\text{COW}(x)\}$—or abbreviating, $\lambda x. \text{COW}(x)$. But one wants to know how the word, which children can acquire, could come to have this semantic value if not because speakers associate ‘cow’ with the concept $\text{COW}(\_)$, and use this concept to construct the corresponding denoter of type $<e_t>$. So I take the psychological hypothesis seriously. On this view, meanings can (qua assembly instructions) exhibit part-whole relations even when the corresponding denotations do not. And CPL, unlike PL, generates expressions that are both significant and subentential: $\lambda x. \text{T}-\text{VALUE}\{\text{COW}(x)\}$ and $\text{BESSIE}$ belong to semantic

---

21 As discussed in chapter six, this proposal seems to be implicit in Heim and Kratzer’s (1998) textbook presentation of a Church/Montague-style semantics for English.
categories, as does \text{T-VALUE}\{\text{COW(BESSIE)}\}. One can similarly treat ‘∼’, which is syncategorematic in PL, as a denoter of type \(<t, t>\) in CPL; and one can introduce a range of denoters for the range of conjunction functions (cp. section three of chapter one). But I haven’t yet said how to assign denotations to quantificational expressions, which are also syncategorematic in PL.

3.4 Functions and Quantification

PL only allows for variables in positions occupied by Fregean singular concepts of type \(<e>\). CPL, as defined thus far, only allows for lambda abstraction into such positions. But one can allow for open sentences like \(\Phi(\text{BESSIE})\) and \(\exists x[\Phi(x)]\), extend the notion of sequence accordingly, and permit abstraction into predicate positions as in \(\lambda \Phi. \text{T-VALUE}\{\exists x[\Phi(x)]\}\). This lambda expression denotes the function \(\exists\) of type \(<et, t>\) such that for each function \(F\) of type \(<e, t>\): \(\exists\) maps \(F\) to \(T\) iff \(F\) maps some entity of type \(<e>\) to \(T\). This higher-order function maps the \text{T-VALUE-function} to \(T\) iff there is a cow.

The new twist adds variables of type \(<e, t>\): \(\Phi, \Phi', \Phi''\), etc. The old variables—\(x, x', x''\), etc.—are of type \(<e>\). The domain already includes functions of type \(<e, t>\). So we can say that \textit{assignments} map the new variables to functions and the old variables of type \(<e>\) to entities. Think of Tarski-sequences, \(n\)-tuples of entities, as special cases of \(n\)-tuples of entity-function \(n\)-tuples. For example, think about replacing \(\langle\text{BESSIE}, \text{VENUS}, \text{SADIE}, \ldots\rangle\) with an \(n\)-tuple, each element of which is a sequence of entity-function pairs with constant entity as indicated below;

\[
\langle\langle\text{BESSIE}, \text{T-VALUE}\{\text{COW}\}\rangle, \langle\text{BESSIE}, \text{T-VALUE}\{\text{HORSE}\}\rangle, \ldots\rangle
\]

\[
\langle\langle\text{VENUS}, \text{T-VALUE}\{\text{COW}\}\rangle, \langle\text{VENUS}, \text{T-VALUE}\{\text{HORSE}\}\rangle, \ldots\rangle
\]

\[
\langle\langle\text{SADIE}, \text{T-VALUE}\{\text{COW}\}\rangle, \langle\text{SADIE}, \text{T-VALUE}\{\text{HORSE}\}\rangle, \ldots\rangle
\]

where each such sequence orders the functions in the same way, allowing for two dimensions of variation in how values are assigned to variables. Let ‘\(A\)’ range over such assignments, each of which can be viewed as a mapping from variable-pairs like \(<x, \Phi>\) to entity-function pairs.

For example, let ALONZO be a particular assignment of the sort indicated above. Then \(\text{ALONZO}(\langle x, \Phi \rangle)\) is the ordered pair \(\langle\text{BESSIE}, \text{T-VALUE}\{\text{COW}\}\rangle\), while \(\text{ALONZO}(\langle x', \Phi'' \rangle)\) is \(\langle\text{VENUS}, \text{T-VALUE}\{\text{HORSE}\}\rangle\). Given how assignments are ordered, we can still speak of the entity that ALONZO associates with the variable \(x\)—viz., Bessie. Each entity \(e\) corresponds to a “horizontal” sequence of entity-function pairs of the form \(<e, F>\). Likewise, the \text{T-VALUE-function} is the function that ALONZO associates with the variable \(\Phi\). Each function \(F\) of type \(<e, t>\) corresponds to a “vertical” sequence of entity-function pairs of the form \(<e, F>\).

The notion of an assignment variant remains the same. If \(A^*\) differs from \(A\) at most with regard to what \(A^*\) assigns to the variable \(x\), then \(A^*\) and \(A\) assign the same value to every other variable, including all the variables of type \(<e, t>\). For example, ALONZO has an \(x\)-variant of the sort indicated below, with Fido replacing Bessie in the \(x\)-row.

\[
\langle\langle\text{FIDO}, \text{T-VALUE}\{\text{COW}\}\rangle, \langle\text{FIDO}, \text{T-VALUE}\{\text{HORSE}\}\rangle, \ldots\rangle
\]

\[
\langle\langle\text{VENUS}, \text{T-VALUE}\{\text{COW}\}\rangle, \langle\text{VENUS}, \text{T-VALUE}\{\text{HORSE}\}\rangle, \ldots\rangle
\]

\[
\langle\langle\text{SADIE}, \text{T-VALUE}\{\text{COW}\}\rangle, \langle\text{SADIE}, \text{T-VALUE}\{\text{HORSE}\}\rangle, \ldots\rangle
\]

Likewise, if \(A^*\) differs from \(A\) at most with regard to what \(A^*\) assigns to the variable \(\Phi\), then \(A^*\) and \(A\) assign the same value to every other variable, including all the variables of type \(<e>\). For example, ALONZO has a \(\Phi\)-variant of the sort indicated below, with the \text{T-VALUE-function} replacing the \text{T-VALUE-function} in the \(\Phi\)-column.
Each entity-function pair corresponds to a truth value in an obvious way. And the notion of satisfaction can be extended accordingly. The open sentence $A(x)$ is satisfied by $A$ if and only if $T$ is the result of applying the function $A(\Phi)$ to the entity $A(x)$.

$$\text{Sat}\{A, \Phi(x)\} \leftrightarrow A(\Phi)(A(x)) = T$$

This extends the original notion of satisfaction via the notion of function application. But the interpretation for the Tarskian existential quantifier remains the same.

$$\text{Sat}\{A, \exists x(\Phi(x))\} \leftrightarrow \text{some } A^*: A^*_x A \text{ is such that } \text{Sat}\{A^*, \Phi(x)\}$$

So the condition imposed by $\exists x(\Phi(x))$ can be specified as follows.

$$\text{Sat}\{A, \exists x(\Phi(x))\} \leftrightarrow \text{some } A^*: A^*_x A \text{ is such that } A^*(\Phi)(A^*(x)) = T$$

If $A^*_x A$, then $A^*(\Phi) = A(\Phi)$. So $A$ satisfies $\exists x(\Phi(x))$ iff some $A^*: A^*_x A$ is such that $A(\Phi)$ maps $A^*(x)$ to $T$. That is, $A$ satisfies $\exists x(\Phi(x))$ iff the function that $A$ assigns to $\Phi$ maps some entity to $T$. In which case, relative to any assignment $A$, $\text{t-value}\{\exists x(\Phi(x))\}$ denotes $T$ or $\perp$, depending on whether or not $A(\Phi)$ maps some entity to $T$. This denotation condition can be specified more formally as follows.

$$\text{Den}\{\text{t-value}\{\exists x(\Phi(x))\}, A\} = T/\perp[\text{Sat}\{A, \exists x(\Phi(x))\}] =$$

$$T/\perp[\text{some } A^*: A^*_x A \text{ is such that } A^*(\Phi)(A^*(x)) = T]$$

So the lambda abstract $\lambda\Phi. \text{t-value}\{\exists x(\Phi(x))\}$ denotes the intended higher-order function.

$$\text{Den}\{\lambda\Phi. \text{t-value}\{\exists x(\Phi(x))\}, A\} = \text{the smallest function } \mathcal{S} \text{ of type } <e, t>$$

such that for each function $F$ of type $<e, t>$, some $A^*: A^*_x A$ is such that $A^*(\Phi) = F$, and $\mathcal{S}$ maps $A^*(\Phi)$ to $\text{Den}\{\text{t-value}\{\exists x(\Phi(x))\}, A^*\}$

Or more briefly, relative to any assignment $A$, $\lambda\Phi. \text{t-value}\{\exists x(\Phi(x))\}$ denotes the function $\mathcal{S}$ that maps each function $F$ to $T$ if and only if $F$ maps some entity to $T$.

An assignment $A$ satisfies $[\Phi'(x) & \Phi(x)]$ iff $A$ satisfies both $\Phi'(x)$ and $\Phi(x)$—i.e., iff both $A(\Phi')$ and $A(\Phi)$ map $A(x)$ to $T$. A function $\mathcal{S}^2$ of type $<e, t, t>$ maps each function $F'$ of type $<e, t>$ to a function $\mathcal{S}$ of type $<e, t>$. So relative to any assignment, the “doubled” lambda abstract $\lambda\Phi'. \lambda\Phi. \text{t-value}\{\exists x[\Phi'(x) \supset \Phi(x)]\}$ denotes the (smallest) function $\mathcal{S}^2$ that maps each function $F'$ to the function $\mathcal{S}$ that maps each function $F$ to $T$ iff some entity $e$ is such that both $F$ and $F'$ map $e$ to $T$. This dyadic higher-order function maps the $\text{cow}$-function to a function that maps the $\text{brown}$-function to $T$ iff some entity is both a cow and brown. Likewise, relative to any assignment, $\lambda\Phi'. \lambda\Phi. \text{t-value}\{\forall x[\Phi'(x) \supset \Phi(x)]\}$ denotes the function $\mathcal{S}^2$ that maps each function $F'$ to the function $\mathcal{S}$ that maps each function $F$ to $T$ iff every entity $e$ is such that $F$ maps $e$ to $T$ if $F'$ maps $e$ to $T$. This dyadic higher-order function maps the $\text{cow}$-function to a function that maps the $\text{brown}$-function to $T$ iff every entity is such that it is brown if it is a cow.
Abstracts like $\lambda \Phi'. \lambda \Phi. \text{t-value} \{ \forall x[\Phi'(x) \supset \Phi(x)] \}$ have Tarskian quantifiers, each binding one variable, as constituents. But since the variable of type $<e>$ is bound, one can abbreviate in a way that highlights the double predicate abstraction: $\lambda \Phi'. \lambda \Phi. \text{every} \{ \Phi', \Phi \}$. Combining $\lambda \Phi'. \lambda \Phi. \text{every} \{ \Phi', \Phi \}$ with $\lambda x. \text{cow}(x)$ yields $\lambda \Phi'. \lambda \Phi. \text{every} \{ \Phi', \Phi \}(\lambda x. \text{cow}(x))$, which can be rewritten as $\lambda \Phi. \text{every} \{ \lambda x. \text{cow}(x), \Phi \}$; and this denoter of type $<e, t>$ can combine with $\lambda x. \text{brown}(x)$ to form $\text{every} \{ \lambda x. \text{cow}(x), \lambda x. \text{brown}(x) \}$, which denotes $T$ or $\bot$ depending on whether or not every cow is brown.

Double predicate abstraction makes it possible to introduce endlessly many concepts—including $\lambda \Phi'. \lambda \Phi. \text{two} \{ \Phi', \Phi \}$, $\lambda \Phi'. \lambda \Phi. \text{three} \{ \Phi', \Phi \}$, etc.—in Fregean fashion. The thought $\text{OneToOne} [\text{cow}(\_), \text{horse}(\_)]$ is true iff the cows correspond one to one with the horses; and the concept $\text{OneToOne} [\_ , \_ ]$ is saturated by two function denoters. So let $\text{OneToOne} [\Phi', \Phi]$ be satisfied by $\mathcal{A}$ iff the (set of) things that $\mathcal{A} (\Phi')$ maps to $T$ correspond one to one with the (set of) things that $\mathcal{A} (\Phi)$ maps to $T$. Then $\lambda \Phi'. \lambda \Phi. \text{OneToOne} [\Phi', \Phi]$ denotes a function $\exists^2$ that maps each function $F'$ to the function $\exists$ that maps each function $F$ to $T$ iff the elements of $\{ x: F'(x) = T \}$ correspond one to one with the elements of $\{ x: F(x) = T \}$. One can likewise introduce $\text{OneToOnePlus} [\_ , \_ ]$ so that $\lambda \Phi'. \lambda \Phi. \text{OneToOnePlus} [\Phi', \Phi]$ denotes a function $\exists^2$ that maps each function $F'$ to the function $\exists$ that maps each function $F$ to $T$ iff: some but not all of the elements of $\{ x: F'(x) = T \}$ correspond one to one with (all) the elements of $\{ x: F(x) = T \}$.

Note that most of the cows are brown iff the brown cows outnumber the cows that are not brown. So given the resources of CPL, $\lambda \Phi'. \lambda \Phi. \text{most} \{ \Phi', \Phi \}$ might be defined in terms of one-to-one correspondence, conjunction, and negation; see note 28 of chapter one.

Thus, CPL supports introduction of many concepts that cannot be defined in PL, many of which humans enjoy; see Rescher (1967), Wiggins (1980). Indeed, CPL supports introduction of function-denoters of every Fregean type, at every level of hierarchy discussed in §1.5 above. The notion of an assignment can be extended to cover variables of any type. And even with regard to monadic predicate abstraction, CPL supports introduction of concepts that go beyond those humans can naturally fetch with words; see chapter five. Staying with quantificational examples here, no word like ‘every’ fetches $\lambda \Phi'. \lambda \Phi. \text{OneToOne} [\Phi', \Phi]$ and combines with two predicates to form a sentence; see Barwise and Cooper (1981), Higginbotham and May (1981). To express this simple second-order concept, we need to form complex locations like ‘The dogs are equinumerous with the cows’. I return to the details and questions raised in chapter six. For now, it is enough to note that CPL allows for concept introduction by exploiting quantification over assignments (abstraction) in ways that may not be natural for humans.

Since the expressions of CPL include endlessly many denoters, and lambda expressions can be used to specify procedures (I-functions as opposed to E-functions), CPL is also suited to the Fregean idea that expressions have Sinnen as well as Bedeutungen. One can say that each denoter of CPL denotes its denotation in a certain way, and that two denoters—$\text{Hesperus}$ and $\text{Phosphorus}$, $\lambda x. |x - 1|$ and $\lambda x. +\sqrt{(x^2 - 2x + 1)}$, $\lambda x. \text{woodchuck}(x)$ and $\lambda x. \text{groundhog}(x)$, etc.—can denote the same entity/function in different ways. So while $\lambda x. \text{woodchuck} (\text{Hesperus})$ and $\lambda x. \text{groundhog} (\text{Phosphorus})$ both denote the truth value $\bot$, one might say that these sentences differ semantically by virtue of denoting in different ways.

---

22 That is, without resorting to first-order specifications like $\exists x[\Phi'(x) \land \Phi(x)] \land \exists x'[\Phi'(x') \land \Phi(x') \land \neg (x = x') \land \forall x''[\Phi(x'') \land \Phi(x'') \supset ((x'' = x) \lor (x'' = x'))]$. Introducing $\lambda \Phi'. \lambda \Phi. \text{seven} \{ \Phi', \Phi \}$ this way is less than ideal.
We needn’t dwell on whether this ends up being a formal variant of an idea noted in section three: WOODCHUCK(HESPERUS) and GROUNDHOG(PHOSPHORUS) differ semantically, even though neither is satisfied by any sequence, because these sentences of PL correspond to different *derivable specifications* of conditions on sequences. The point here is that a Tarskian semantics for PL can be extended in a way that provides at least much of what Frege wanted. This provides a potential model for human languages. Though again, CPL puts quantification over sequence variants to work in a way that permits representation of any computable function. And this may be more work of a kind that human minds cannot do in any natural way.

5. Summary and Transition

Thanks to Frege, we have modern quantificational logic and a vivid case study of how a language can be used to introduce formally new concepts that are useful for certain cognitive purposes. Thanks to Tarski, we have a paradigm of what it is for a language to have a truth conditional semantics. Thanks to Church, we have a clear example of how such a language can be extended in interesting ways. This invites an exciting project to which many theorists have contributed a great deal of insightful work: try to successively modify PL and/or CPL in ways that yield increasingly better models of human languages. My own proposals are anchored within this tradition. In my view, humans enjoy a language of thought that can be viewed as an extension of a fragment of PL: PL does not allow various constructions that human languages permit; though in other respects, PL (and hence, CPL) is far more permissive. But there are at least two ways—sketched in this final section—to think about modifying a language designed to have a truth conditional semantics, in the service of providing better models of human languages.

One possibility, urged by Davidson (1967b, 1984) and Montague (1974), is that each human language *has* a truth conditional semantics. From this perspective, modifications of (C)PL can be viewed as attempts to characterize the satisfaction/denotation conditions that human linguistic expressions impose on the entities/functions that speakers can talk about and quantify over. As discussed in the introductory chapter, (D1) is a compact formulation of this hypothesis.

(D1) for each human language H, there is a correct theory of truth for H.

And if (D1) is plausible, or at least not obviously wrong, then (D2) can be tempting.

(D2) a suitably formulated theory of truth for a human language H can serve as the core of a correct theory of meaning for H

The alternative urged here is that human languages let their users introduce, combine, and express concepts that are useful but not ideal in Frege’s sense.

It is an old idea that serious inquiry includes and may require attempts to *construct* languages that mirror reality, and to understand how actual modes of thought and talk *diverge* from this scientific ideal.23 From this perspective, it would take a miracle—a universe ordered so that human biology naturally reflects it—for a human I-language to have a truth conditional semantics. One can view sentences of (C)PL as models of the thoughts that humans can form by using *all* the resources, natural and theoretical, at our disposal. But if we want models of the thoughts that we regularly form in a natural way, then we may need to invent model languages—perhaps extensions of fragments of PL—with *that* purpose in mind. We may also need to think of spoken/signed languages as tools used in the assembly of natural thoughts, rather than mere systems for communicating thoughts constructed in some other way.

23 Descartes tried to think with clear and distinct ideas governed by the *Principles of Philosophy*. Leibniz wanted a *Characteristica Universalis*. Following Frege and Russell, Wittgenstein envisioned the possibility of a language that has a truth conditional semantics as outlined in the *Tractatus*. 
Perhaps (1) is like the ideal sentence \textsc{precedes}(\textsc{zeros}, \textsc{one}) in having a truth condition.

(1) Snow is white.

Theorists can hypothesize that for some condition \( C \) on assignments of values to variables in (1), snow is white iff for each assignment \( \mathcal{A} \), \( C(\mathcal{A}) \). But is this a correct hypothesis about a human linguistic expressions, or an idealization whose limitations need to be highlighted in an account of what meanings are? Should we grant that (1) is true iff snow is white, and try to figure out how this could be so, or should we treat biconditionals like (2)

(2) ‘Snow is white.’ is true iff snow is white.

as initially useful but ultimately distorting simplifications? In the next two chapters, I offer reasons for not following the first path; in my view, (D1) and (D2) are both implausible. In chapters five and six, I return to my proposed alternative.
Chapter Three: Event Variables, Framing, and Logical Form

The invented languages described in chapter two were designed with logic and arithmetic in mind. And while human languages can be used to make claims about numbers—e.g., that zero precedes one—they are more often used to report on the passing show. Moreover, such reports can contain boundlessly many adverbial modifiers that initially seem to have no analogs in the Tarskian language PL or its Church-style extension. So at a minimum, one wants to see how (D) (D) for each human language H, there is a Tarski-style theory of truth that can serve as core of an adequate theory of meaning for H can accommodate sentences like (1) and (2).

(1) Al chased Theo.
(2) Al chased Theo gleefully.

According to Davidson (1967a), action reports have truth conditions that can be specified with existential generalizations as in (1a) and (2a).

(1a) ∃x[CHASED(X, AL, Theo)]
(2a) ∃x[CHASED(X, AL, Theo) & GLEEFUL(X)]

At first, this proposal seems to support thesis (D). But upon reflection, examples like (1-6)

(3) Theo chased Al.
(4) Theo chased Al gleefully.
(5) Theo chased Al gleefully.
(6) Al chased Theo gleefully.

show that simple cases of adverbial modification still present serious difficulties for (D1).

(D1) for each human language H, there is a correct theory of truth for H

This does not establish that (D) is false. But the facts concerning adverbial modification do not merely tell against (D), they undercut much of the motivation for defending it.

1. Reporting What Happened

One might wonder how any plausible truth conditions for (1) and (2) could be specified compositionally. Davidson offered an insightful answer. But like many insights, it was packaged in a larger theory whose initial attractions fade. In this first section, I provide an overview of the basic problem, and discuss potential responses in subsequent sections.

1.1 Not Enough Events

Given a language that generates tensed expressions like (1), but no examples of adverbial modification like (2), theorists might hope to specify the meaning of (1) with (1b).

(1b) ∃x:BEFORE(x, T)[CHASES-AT(AL, Theo, x)]

An assignment A satisfies CHASES-AT(AL, Theo, x) iff A(x) is a time at which A(AL) tenselessly chases A(Theo); AL and Theo are constants, variables whose values do not vary; T is an index for moments/intervals of time; and a pragmatically relevant assignment assigns the relevant moment/interval to T. This encodes the idea that ‘chase’ is true of ordered triples <α, β, γ> such that α chases β at γ. On this view, the verb in (1) combines with three arguments, the first of which is a tense morpheme that can be viewed a restricted quantifier akin to ‘sometime before now’. Though for now, let’s not worry about the difference between (1b) and (1b′).

(1b′) ∃x:[BEFORE(x, T) & CHASES-AT(AL, Theo, x)]

A slightly different option is (1c);

(1c) ∃x:BEFORE(x, T)[AT-CHASES(X, AL, Theo)]

where A satisfies AT-CHASES(X, AL, Theo) iff A(x) is a time at which A(AL) tenselessly chases
$\mathcal{A}(\text{Theo})$. More briefly, $\mathcal{A}(\text{chases}(x, \text{Al}, \text{Theo}))$ is true of times at which Al chases Theo. With regard to logic, this is notational variation. But order of pronunciation may not reflect order of grammatical arguments with regard to tense. And one might hope to analyze (7) as in (7a),

(7) Al chased Theo yesterday.

(7a) $\exists x: \text{BEFORE}(x, t) [\text{chases}(x, \text{Al}, \text{Theo}) \land \text{yesterday}(x)]$

with $\text{yesterday}(x)$ imposing a second temporal restriction.\footnote{Let’s grant that ‘yesterday’ can be used to talk about a proper part of the day before the day of speech. Let’s also bracket the difficulties of specifying an extension for ‘day’—given calendar changes, cesium clocks, and leap seconds (see King 200x)—and words like ‘Monday’, which can appear in ‘on Monday’ or ‘a Monday’.} Still, this doesn’t accommodate (2).

(2) Al chased Theo gleefully.

A committed theorist might insist that times can be gleeful, and that (2) is true iff (2b) is.

(2b) $\exists x: \text{BEFORE}(x, t) [\text{chases}(x, \text{Al}, \text{Theo}) \land \text{gleeful}(x)]$

But the mere existence of gleeful times, whatever they are, doesn’t make this analysis plausible. Suppose that a certain minute—say from 12:00 to 12:01 (EST) on Dec. 25, 1970—was the only time at which Al chased Theo, and also the only time at which Dave chased Simon. If (2) and (8) have truth conditions, then presumably, both can be true in this scenario.

(8) Dave chased Simon gleefully.

Al may have chased Theo gleefully, while Dave chased Simon gleefully. But the minute in question was not both gleeful and gleeeless. Moreover, Al may have chased Theo athletically though not skillfully, while Dave chased Simon skillfully though not athletically; and so on.

Nonetheless, one might think that a small modification of (2b) yields major improvement. We needn’t say that an assignment $\mathcal{A}$ satisfies $\text{BEFORE}(x, t)$ only if $\mathcal{A}(x)$ is a time; an event can occur before a given time. And instead of trying to specify the meaning of ‘chase’ in terms of $\text{times at which}$ one thing chases another, one can appeal to $\text{events in which}$ one thing chases another. So following Davidson, one might start with (2a) and elaborate it as in (2a');

(2a) $\exists x [\text{chases}(x, \text{Al}, \text{Theo}) \land \text{gleeful}(x)]$

(2a') $\exists x: \text{BEFORE}(x, t) [\text{chases}(x, \text{Al}, \text{Theo}) \land \text{gleeful}(x)]$

where $\mathcal{A}$ satisfies $\text{chases}(x, \text{Al}, \text{Theo})$ iff $\mathcal{A}(x)$ is a chase by $\mathcal{A}(\text{Al})$ of $\mathcal{A}(\text{Theo})$. This raises the question of what it is for events to be gleeful. But at least with regard to a chase, one might say that it is gleeful (in the relevant sense) if it is done gleefully, and that it is gleeful if done gleelessly. We can also replace (7a) with (7b), construing $\text{yesterday}(x)$ accordingly; cp. note 1.

(7b) $\exists x: \text{BEFORE}(x, t) [\text{chases}(x, \text{Al}, \text{Theo}) \land \text{yesterday}(x)]$

At first, this looks like good news for thesis (D).

(D) for each human language $\mathbf{H}$, there is a Tarski-style theory of truth that can serve as core of an adequate theory of meaning for $\mathbf{H}$

Davidson showed how a finitely specifiable theory could assign plausible truth conditions, in a way that captures apparent implications, to boundlessly many sentences that contain adverbial modifiers. This also illustrated a potentially more general strategy for dealing with apparent counterexamples to (D): specify meanings in terms of variables that range over things that one might not initially think of, especially if one was initially thinking about arithmetic. But the mere existence of gleeful events and gleeeless events, whatever they are, isn’t enough to make the Davidsonian analyses plausible. For the goal is not merely to accommodate (2) and (8).

(2) Al chased Theo gleefully.

(8) Dave chased Simon gleelessly.
These sentences have different parts. We want a compositional theory of meaning that plausibly accommodates (2) and (5), (4) and (6), etc.

(5) Theo chased Al gleefully.
(4) Theo chased Al gleelessly.
(6) Al chased Theo gleelessly.

No one of these sentences, taken on its own, may seem especially problematic. And even taken together, it is easy to focus on relatively unproblematic scenarios in which Theo chased Al on some occasion that differs from the occasion on which Al chased Theo. But suppose that Theo chased Al exactly once, in the very same minute that Al chased Theo. Put another way, suppose that Al and Theo chased each other. Neither chipmunk caught up on the other before they both gave up. Al chased Theo gleefully and athletically, but not skillfully; Theo chased Al gleelessly and unathetically, but skillfully. Given this scenario, we can correctly report some of what happened with (2) and (4), though not with (5) or (6). Likewise, we can correctly report some of what happened with (9) and (10), though not with (11) or (12).

(9) Al chased Theo gleefully and athletically but not skillfully.
(10) Theo chased Al gleelessly and unathetically but skillfully.
(11) Al chased Theo gleelessly or unathetically or skillfully.
(12) Theo chased Al gleefully or athletically or not skillfully.

Davidsonian analysis thus seems to require that the envisioned scenario include at least two chases: a chase by Al of Theo that was gleeful, athletic, and not skillful; and a chase by Theo of Al that was gleeless, unathletic, and skillful. But prima facie, there were not two chases with the same participants in the same region of spacetime. The event of Al moving, during that minute, can be described as both an event of him (intentionally) running after Theo and an event of him (unintentionally) maintaining a constant distance from Theo; and likewise, mutatis mutandis, for the simultaneous but distinct event of Theo moving. So the net effect can be described as Al chasing Theo, or as Theo chasing Al. But in my view, there was only one chase, in which both chipmunks participated. No solo-chipmunk episode was an event of chasing Theo or chasing Al; in a chase, the things chase is a moving participant. We can speak of two chases along one path. Dave may have chased Simon around a table, around which a dog chased a cat. But if Al and Theo chased each other around a tree, around which nothing else was ever chased, there were not two chases around that tree. Counting chases individualistically seems wrong.

Similarly, in my view, no event in which Miss Scarlet is the sole participant is an event of her stabbing Colonel Mustard with a knife. In any such event, the colonel and the knife are involved. And the general point is not restricted to cases of intentional action. A red ball can collide with a green one; but the collision, which has no direction or momentum, is not the motion of either ball. I’ll return to all these examples. Though in the end, it doesn’t matter which events get called chases/stabbings/collisions for purposes of describing the world as best we can. Say if you like that there were two chases around the tree, one per chipmunk. Then the question is whether chases, so counted, are plausibly things that ‘chase’ is true of. I think the answer is negative: given event analyses, ‘chase’ is not true of anything in which no chasee is chased.

1.2 Potential Responses and the Bigger Point

One might say that a chase by Al of Theo—or at least an event of Al chasing Theo—need not be a chase. Even if the solo-chipmunk events are not chases, they might be events e1 and e2 such that: ‘chased’ is true of both <e1, Al, Theo> and <e2, Theo, Al>; e1 is done by Al, gleeful, and so on; e2 is done by Theo, gleeless, and so on. Perhaps e1, Al, and Theo are (so-ordered) related chase-wise in part because e1 co-occurred with a suitable event of Theo moving. But even if this
maneuver works for some cases, it doesn’t help with Miss Scarlet’s knife. If some event, Scarlet, and Mustard are (so-ordered) related stab-wise, that is presumably because of the knife she used in that event. I also think that stressing relationality makes a mystery of some simple implications. For example, (1) apparently implies that Theo was chased,

(1) Al chased Theo.

as if a chase by Al of Theo is an event in which Theo gets chased. Moreover, we do talk of chases, and not as if they are ordered triples in which chasees “participate” abstractly. Indeed, we use the monadic notion when explicating \( \text{CHASE}(X, X', X'') \) as \( \text{CHASE}(X) \& BY(X, X') \& \text{OF}(X, X'') \); cp. Castañeda (1967). But when we talk of chases, we describe “what happened” as if the transpiring events have multiple participants.

That said, human languages also associate grammatical positions with thematic relations, as if events themselves exhibited subject-object asymmetries. We speak as though Al was the agent in his chase of Theo, who was the agent in his chase of Al. Yet we also speak of the chase as an episode, or perhaps some episodes, in which each chipmunk ran. And in the right contexts, we can use ‘chase’ to talk about Al’s attempt to catch Theo.2 But this doesn’t show that ‘chase’ is true of any such events. On the contrary, event analyses suggest that if ‘chase’ is true of any events, then these events have chasers (agents) and chasees (patients) as participants. I don’t deny that event analyses also suggest that if ‘chase’ is true of any events, then these events do not have their patients as participants. Though in my view, the right conclusion is that ‘chase’ does not have a satisfaction condition. It has a meaning without being true of anything.

In this respect, I think talk of chases is like talk of sunrises, skies, and other “things” we talk about without using words that are true of them. As noted in chapter zero, (D1) seems especially implausible with regard to words like ‘sky’,

(D1) for each human language \( H \), there is a correct theory of truth for \( H \) making thesis (D) is correspondingly implausible.

(D) for each human language \( H \), there is a Tarski-style theory of truth that can serve as core of an adequate theory of meaning for \( H \)

Even if some cleverly constructed truth theory can accommodate ‘sky’ and ‘night’, such a theory is unlikely to capture the meaning of ‘night sky’. It can be tempting to set aside such words aside, along with ‘sunrise’ and ‘Venice’, as special cases that introduce special complications. But I don’t think the objections to (D) can be quarantined. In this chapter, I try to make the point by arguing that while event analyses are motivated beyond reasonable doubt, such analyses lead to the conclusion that verbs like ‘chase’ (‘stab’, ‘collide’, ‘strike’, ‘pull’, ‘play’, ‘see’, ‘hear’) do not have satisfaction conditions. The \textit{verbs} in ‘saw Bellerophon chase Pegasus’ tell against (D).

In chapter four, I argue that quite apart from considerations regarding events, (D) leads to paradox in ways that make it independently implausible to suppose that sentences like (1-12) have truth conditions. I don’t, however, pretend to offer \textit{proofs} that (D) is false.

Since (D) is a hypothesis, regarding the I-languages that human children naturally acquire, we shouldn’t expect proofs. My aim is to argue that (D) is less plausible than the claim that meanings are instructions for how to build concepts. But precisely because (D) is a \textit{hypothesis}, we must guard against conflating “regimentations” of human languages—proposals about how we ought to speak, at least when theorizing, given the available values for Tarski-style variables—with claims about the actual meanings of human linguistic expressions.

---

2 I suspect that theorists can use ‘action of chasing Theo’ to talk about a mental event of trying that can be a cause of certain bodily motions. See Pietroski (1998), drawing on Hornsby (1980) and Thomson (1977) among others.
In particular, it is recurrently tempting to say that some event of Al moving was a chase of Theo (or at least an event e1 such that ‘chase’ is true of \(<e1, Al, Theo>\)) by virtue of being suitably related to some event of Theo moving, which was itself a chase of Al by virtue of being suitably related to the event of Al moving. One can imagine talking this way, so that there were two chases in the one minute; though at each moment m, each of these solo-chipmunk events has a chipmunk in a different place. But this hardly shows that the lexical item ‘chase’, which human children acquire with ease, is true of these one-participant events. Likewise, I don’t deny that we can speak of the chase in which Al and Theo both participated, or at least the events that together constituted them chasing each other. I just deny that ‘chase’ is true of these multi-participant events. Though to repeat, the real issue is not about what gets to be called a chase.

One can stipulate that \(\text{CHASE}(X, AL, THEO)\) and \(\text{CHASE}(X, THEO, AL)\) have different satisfiers, and one can insist that the relevant values of the invented variable are two-participant events, exactly one of which is gleeful. But this option turns out to be especially unpalatable if the goal is to defend thesis (D). Since the requisite “events” seem to be individuated about as finely as their linguistic descriptions, one challenge is to make it plausible that each of them can be described in boundlessly many ways—e.g., as a chase of Theo by Al, a gleeful thing, and so on—as required by Davidsonian analyses. Advocates of (D) need to individuate values of posited variables in a way that is plausible given a truth-theoretic conception of meaning.

In particular, it is bad news for (D) if defending it with regard to sentences like (2)

\[
(2) \text{Al chased Theo gleefully.}
\]

already requires appeal to domain entities that seem to individuated in terms of meanings. Put another way, it’s not enough to show that (D1) is defensible

\[
(\text{D1}) \text{ for each human language } H, \text{ there is a correct theory of truth for } H.
\]

given some suitably generous metaphysics. The interesting proposal was that we should embrace (D), despite apparent objections, because (D1) turns out to be more plausible than expected; see §3 of chapter zero. Sentences like (2) were supposed to be parade cases that motivated (D1), not apparent counterexamples that might be handled given sufficient ingenuity. Davidson and others thus tried to defend a conception of events as relatively “coarse grained” entities that are independent of how we think and talk about them. (The events were supposed to be there, along with the chipmunks, as things speakers can quantify over and describe in many ways. And we can agree to bracket questions about how to individuate chipmunks, at least until someone posits two of them in the same place at the same time. Proliferating chases spoils the project.)

In reply to some examples, relativizing to descriptions may be legitimate. In particular, both (13) and (14) might be used to describe “something Simon did” on a certain occasion.

\[
(13) \text{Simon swam the channel quickly.}
\]

\[
(14) \text{Simon crossed the channel slowly.}
\]

This is compatible with a version of (D1) that posits just one event of Simon getting from one side of the channel to the other: the one event can be quick for a swimming of that distance, yet slow by usual standards for a channel-crossing. But as we’ll see in section four, this strategy of “relativization” does not plausibly extend to examples like (1-12).

Spelling this out requires pages, given the possibility of clever Davidsonian replies and responses to objections to those replies. But in the end, I think the moral is simple: analyzing (2) as (2a), elaborated as (2a’), was a good idea;

\[
(2a) \exists x[\text{CHASE}(X, AL, THEO) \& \text{GLEEFUL}(x)]
\]

\[
(2a’) \exists x: \text{BEFORE}(x, T)[\text{CHASE}(X, AL, THEO) \& \text{GLEEFUL}(x)]
\]
but quantifying over events, as opposed to times, merely delays the objections to (D1). Events are individuated more finely than times, but not finely enough.\(^3\)

This moral is largely independent of my concerns about specifying human linguistic meanings with Tarskian ampersands, quantifiers, and invented sentences like \(\text{CHASED}(x, x', x'')\) and \(\text{GAVE}(x, x', x'')\) that have more than two variables.\(^4\) So in this chapter, I won’t object to such apparatus. I do, however, think that the virtues of specifying adverbial meanings conjunctively are best preserved by recasting Davidson’s proposal as an illustration of how phrasal meanings can be fruitfully viewed as instructions for how to build conjunctive monadic concepts.

### 1.3 Form vs. Ontology

Let me stress my desire to maintain the standard practice of using intuitions about the impeccability of certain inferences as data for theories of meaning. Competent speakers find certain inferences compelling. Sometimes, this is due in part to background beliefs that can be represented as logically contingent premises. But I assume that many inferences are compelling because they exhibit patterns that speakers recognize as having a special formal character. Indeed, I want to protect this assumption from the idea that compellingness is to be diagnosed in terms of truth preservation as opposed to natural forms of the sort described in chapter one.

One can argue about cases and details. But to repeat a standard example, instances of ‘\(\Phi\) is a brown \(\Phi\), so \(\Phi\) is a \(\Phi\)’—or at least the thoughts expressed via such instances—seem impeccable. If what you would say with (15) is true, then what you would say with (16) is true.

\[
\begin{align*}
(15) & \quad \text{Simon is a brown chipmunk}. \\
(16) & \quad \text{Simon is a chipmunk}.
\end{align*}
\]

Put another way, speakers of English know that (16) can be used to tell the truth in any context where (15) can be used to tell the truth. However we formulate the generalization, the phenomenon seems to be that speakers understand simple adjectival modification as in ‘brown chipmunk’ as a sign of predicate conjunction, and they recognize conjunct reduction as a form of impeccable inference in “default” settings like (15); cp. ‘not a brown chipmunk’.

Davidson rightly noted that the inference from (2) to (1) is equally compelling,

\[
\begin{align*}
(2) & \quad \text{Al chased Theo gleefully}. \\
(1) & \quad \text{Al chased Theo}.
\end{align*}
\]

and that this suggests his proposed logical forms.

\[
\begin{align*}
(2a) & \quad \exists x[\text{CHASED}(x, \text{AL}, \text{THEO}) \& \text{GLEEFUL}(x)] \\
(1a) & \quad \exists x[\text{CHASED}(x, \text{AL}, \text{THEO})]
\end{align*}
\]

But we can detach this good idea from the implausible claim that (2) has a truth condition that is specified by (2a). We can instead view (2a) as a Tarskian analog of a human thought that can be assembled by executing instruction (2); where the human thought may be more like an existential closure of the monadic concept indicated below.

\[
\exists[\text{AGENT(_,_)\&\text{AL(_)}\&\text{CHASE(_)}\&\text{PAST(_)}\&\exists[\text{PATIENT(_,_)\&\text{THEO(_)}]\&\text{GLEEFUL(_)}
\]

If our use of such concepts is governed by a natural logic that permits conjunct-reduction in this setting, then the data point that (2) implies (1) can be captured without supposing that sentences of a human I-language have truth conditions.

\[\text{---}\]

\(^3\) Analogously, the sets of possible worlds may vastly outnumber the truth values. Hence, many objections to identifying sentence meanings with truth values are not objections to identifying these meanings with sets of worlds. Yet boundlessly many examples suggest that sentence meanings are individuated more finely than sets of worlds.

\(^4\) And as we’ll see, replacing \(\text{CHASE}(x, \text{AL}, \text{THEO})\) with \(\text{AGENT}(x, \text{AL}) \& \text{CHASE}(x, \text{THEO}) \& \text{PATIENT}(x, \text{THEO})\)—thereby avoiding appeal to expressions with more than two variables—only sharpens the objections to (D).
Instead of treating “logical forms” like (2a) as specifications of truth conditions that sentence like (2) have, we can view the invented sentences as models of thoughts that might be constructed by a mind that executes certain concept-assembly instructions. From this perspective, the forms are what matter, not the stipulated truth conditions. I’ll say more about utterances and truth in chapter four. But even if sentences like (1) and (2) do not have truth conditions, many utterances of such sentences can be true.

Speakers often use expressions to correctly describe (some of) what happened. One might tell the truth by saying that Al and Theo chased each other, or that two chipmunks ran around a tree. Witnesses often report what happened in different correct ways. But this doesn’t show that some assignment $A$ satisfies both $\text{CHASE}(X, \text{AL}, \text{THEO})$ and $\text{CHASE}(X, \text{THEO}, \text{AL})$. Expressions may be used to assemble concepts that can be used in contextualized acts of judgment/assertion that are often verified by what happened. If sentences of a human language do not have truth conditions, they do not have truth conditions that are specified by other sentences. In which case, (2) does not bequeath truth conditions to utterances of it. But that’s OK: truth-evaluable episodes don’t need to have truth conditions, much less inherit them from sentences.

A particular act of judgment/assertion might be so perspicuous that for each way the world could be, the act determines whether it would be true or false if the world were that way. For some purposes, we at least aspire to such perspicuity. But often, we settle for less. An act of judgment/assertion may be clear enough to be true or false, given how the world is, in part because the relevant context presents an easy call regarding the question at hand; see §3.1 of chapter zero, §5.1 of chapter four, and the history of contract law. Moreover even if there is a function from contexts of using a sentence $S$ assertively to truth values, it hardly follows that $S$ “has” the function in the sense that $S$ has a meaning. So we shouldn’t assume that thesis (L) is true, just because sentential utterances are often true.

$$(L) \text{ if } S \text{ is a sentence of a human language, and } S^* \text{ is the logical form of } S,$$
$$\text{then } S \text{ has the truth condition that } S^* \text{ specifies}$$

We can identify logical forms with invented sentences that are designed to have truth conditions, and then use the inventions to model human language sentences for certain purposes. But (L) remains a conjecture, not a truism. The sentences we naturally use have meanings that determine logical forms. But these sentences may have meanings without having truth conditions.

That said, if sentences like (2) don’t have truth conditions,

(2) Al chased Theo gleefully.

one wants to know how humans (or at least philosophers) get tempted into making incompatible claims/judgments about what happened. One wants some way of thinking about the puzzles in this vicinity—often described in terms of events and how they are individuated—that does not demand a resolution in the form of a truth-theoretic semantics for English (or any other human language) and a substantive proposal regarding values of the alleged variables. So before getting back to linguistic details in section three, let me offer a reminder that thinkers are often susceptible to “framing effects,” some of which run deep. My suspicion is that we can easily slide into inconsistency if we think about an activity that involves more than one agent by using a concept that was constructed by executing an instruction like (2), whose grammatical “subject” is singular. But especially in light of Kahneman and Tversky’s work, we shouldn’t be surprised if using certain linguistic forms in certain contexts leads us logically astray.

---

5 At least relative to a choice of a suitably expressive metalanguage $M$. Though given indeterminacy, “the” logical form of $S$ relative to $M$ may be a class of $M$-sentences; see Quine (1960), Davidson (1984).
I think that linguistic framing is a regular feature of normal cognition for speakers of a human language. We automatically think about things (in the intentional sense of ‘think about’) by using concepts that we have assembled by using linguistic expressions, which have their own formal character; and often, the resulting concepts are not true of representation-independent things. Once in a while, we even notice this. In my view, the right response to this aspect of the human condition is to deny that (2) has a truth condition, instead of insisting that the world is the way it needs to be for sentences like (2) to have truth conditions.

2. Getting Framed

Kahneman (2011) helpfully reviews the literature on various kinds of framing effects, including those that he and Tversky made famous; see Kahneman & Tversky (1979, 2000), Tversky & Kahneman (1981), Kahneman et.al. (1982). Some of the examples reveal that in many respects, we humans are lazy thinkers who are apt to answer questions in ways that initially sound good, even if a moment’s reflection would show the answer to be wrong. In §2.2, I discuss a more troubling example (due to Thomas Schelling) as a potential analogy to “event framing.”

2.1 Warm Up Effects

Consider the famous question in (17).

(17) A bat and a ball cost $1.10. The bat costs a dollar more than the ball. How much does the ball cost?

Even among undergraduates at elite institutions, more than half answer that the ball costs ten cents. But then the bat would cost a dollar. And a dollar is not a dollar more than ten cents.

Finding the right answer is not conceptually difficult. Though it does require a little concentration: \(x + y = 1.1, \) and \(x = 1 + y; \) so \(1 + 2y = 1.1; \) so \(y = .05, \) and the ball costs a nickel. Given this, one might have expected some correct answers, some calculation errors, and various indirect responses like ‘I don’t know’, ‘Less than a dime’, ‘Who cares’, or something less polite. But the majority answer is not just wrong. It is obviously wrong if you think about it all. Now in some sense, it isn’t news that people often answer questions without thinking about their answers. We all know that often, we don’t ask ourselves whether what initially sounds good is at all plausible. Though the implications of this point are especially depressing, given further evidence about the factors that can make one answer initially sound better than another.

In the bat and ball case, it seems that we are tempted to answer a simpler question—how much is left if you take a dollar away from a dollar and a dime?—instead of the question posed. In other cases, we are apparently influenced by differences in the affective responses associated with provably equivalent descriptions of a situation or choice. Even for surgeons, a ninety-five percent survival rate sounds better than a five percent fatality rate. Such cases provide classic examples of framing, since one can show the effect of (re)presenting a situation/choice in one way rather than another. But the sources of framing need not be affective, at least not primarily.

Consider another example from Kahneman and Tversky. Adam and Beth drive equal distances each year, and each just bought a new car. Adam was getting twelve miles per gallon, but now he gets fourteen. Beth was getting thirty miles per gallon, but now she gets forty. Who will save more gas? There is a powerful inclination to say that Beth will save more. Yet we know how to determine the answer: pick a distance D; calculate \(D/12, D/14, \) and subtract to determine how many gallons of gas Adam will save; calculate \(D/30, D/40, \) and subtract to determine how many gallons of gas Beth will save. Most of us also know that we are not good at dividing by fourteen in our heads. Yet the human inclination is not to say, “I do not know; so if the question is worth answering at all, I shall get a pencil and work it out.”
Rather, we find ourselves inclined to blurt out “Beth will save more”—evidently because the difference between forty and thirty is bigger, both absolutely and as a ratio, than the difference between fourteen and twelve. This is, however, not a good reason for concluding that Beth will save more gas than Adam. If they both drive 10,000 miles in a year, Adam will save about 119 gallons, while Beth saves about 83. (Interestingly, if the question is framed in terms of gallons per mile, we are less likely to think that we know the answer.)

These effects are fascinating, in part because we can easily see that our initial inclinations are not only wrong, they are objectively inferior to other responses that are available and justifiable via reasoning that we are fully capable of conducting. Correlatively, discovery of framing effects can have policy implications. But in some cases, framing effects run deeper, making them intellectually interesting in another way—akin to paradoxes—with policy implications that are less clear, apart from suggesting epistemic modesty. Kahneman (2011) reports that his favorite example of this sort is one that (Nobel Prize winner) Thomas Schelling used in the classroom. This beautiful illustration concerns tax law. But I think it also provides a useful analogy to the puzzles posed by sentences like (9) and (10).

(9) Al chased Theo gleefully and athletically but not skillfully.
(10) Theo chased Al gleefully and unathletically but skillfully.

2.2 Deep Framing

Schelling asked his students to think about the policy of reducing taxes for those who have (dependent) children. Suppose your income tax depends entirely on your (household) income and how many children you have. For each income $i$ and number $k$ of children, there is a tax $t$:

$$\text{Tax}(i, k) = t.$$  

The “child deduction” might be flat, say a thousand dollars per child. That is, each income can be paired with a “base” tax, from which some multiple of 1000 is subtracted:

$$\text{Tax}(i, k) = \text{Base}(i) - k[1000].$$  

Alternatively, one might adopt a system in which the deduction for each child depends on household income:

$$\text{Tax}(i, k) = \text{Base}(i) - k[\text{Deduction}(i)].$$  

Given these options, there are many policy questions. But consider (8), which seems relatively easy.

(18) Should the child deduction be larger for the rich than for the poor?

At least for many of us, it seems unfair to adopt a graduated deduction policy, and then make the deduction per child larger for those who already have larger incomes. Hold that thought.

By thinking in terms of deductions, we effectively take the “standard household” to be childless. The base tax is what a childless household pays. But we could instead assume two children per household, start with a lower base tax for all incomes, and impose a surcharge on households with fewer than two children (e.g., $1000 for each child less than two):

$$\text{Tax}(i, k) = \text{Base}^*(i) + (2 - k)[1000];$$ where for each income $i$, $\text{Base}^*(i) = \text{Base}(i) - 2000$. We could also let the surcharge depend on income:

$$\text{Tax}(i, k) = \text{Base}^*(i) + (2 - k)[\text{Surcharge}(i)];$$ where $\text{Base}^*(i) = \text{Base}(i) - 2[\text{Deduction}(i)].$ Again, this presents various questions. But consider (19).

(19) Should the childless poor pay as large a surcharge as the childless rich?

Given a system that penalizes childlessness, with higher taxes for each income, at least many of us would find it unfair to make the poor pay as large a penalty as the rich. A childless poor household would sacrifice a greater percentage of income, for being childless, than a childless rich household. So one wants to say that any such surcharge should be graduated, with the childless poor paying a smaller surcharge than the childless rich. But if you answered both (18) and (19) negatively, then you endorsed a contradiction.

---

For simplicity, assume that no household has more than two children. But it doesn’t matter if there is also a tax deduction for each child beyond the second, or if we take the standard household to have ten kids (reduce the base tax, and impose surcharges accordingly). Though in principle, the last policy might promote population growth.
As Kahneman puts the point, for any given income, the difference between the tax owed by a two-child family and by a childless family can be described as a reduction or as an increase. And if poor households are to receive at least the same benefit as the rich for having children, then poor households must pay at least the same penalty as the rich for being childless. In the abstract, this seems obvious. Still, it can be remarkably hard to shake the sense that both (18) and (19) deserve negative answers. I had to stare, for a long time, at a proof that (20) is not possible.

\begin{equation}
(20) \sim [\text{Deduction}(i_{\text{high}}) > \text{Deduction}(i_{\text{low}})] \& [\text{Surcharge}(i_{\text{low}}) < \text{Surcharge}(i_{\text{high}})]
\end{equation}

For each income, high or low, the deduction has to be the same as the surcharge. One family’s deduction is another family’s surcharge. So (21) and (22) are obviously true.  

\begin{align}
(21) & \text{Deduction}(i_{\text{high}}) = \text{Surcharge}(i_{\text{high}}) \\
(22) & \text{Deduction}(i_{\text{low}}) = \text{Surcharge}(i_{\text{low}})
\end{align}

Given (21), the second conjunct of (20) implies (23), which might seem fine by itself.

\begin{equation}
(23) \text{Surcharge}(i_{\text{low}}) < \text{Deduction}(i_{\text{high}})
\end{equation}

But (23) and (22) imply (24), which is incompatible with the first conjunct of (22).

\begin{equation}
(24) \text{Deduction}(i_{\text{low}}) < \text{Deduction}(i_{\text{high}})
\end{equation}

The inferences are uncomplicated: if \( \alpha = \beta \), and \( \gamma < \beta \), then \( \gamma < \alpha \); if \( \alpha < \beta \), and \( \gamma = \alpha \), then \( \gamma < \beta \). And yet, our—or least my—gut responses to (18) and (19) remain. Quite humbling, and worth remembering when one uses the Fregean toolkit for logic to characterize linguistic meanings.

One might conclude that since (18) clearly deserves a negative answer, we must answer (19) affirmatively. But even if one answers (25) affirmatively, after thinking about (19-24),

\begin{equation}
(25) \text{Should the child deduction be flat?}
\end{equation}

it still seems that (26) should be answered negatively.

\begin{equation}
(26) \text{Should there be a flat tax on childlessness?}
\end{equation}

We could eliminate the child deduction. But given a flat deduction, poor households with kids get more tax relief as a percentage of income than rich households. That raises question (27).

\begin{equation}
(27) \text{Should we eliminate a tax break for poor families with children?}
\end{equation}

Kahneman (2011) draws a dramatic and disturbing conclusion.

The message about the nature of framing is stark: framing should not be viewed as an intervention that masks or distorts an underlying preference. At least in this instance...there is no underlying preference that is masked or distorted by the frame. Our preferences are about framed problems, and our moral intuitions are about descriptions, not substance (pp. 412-413)?

I take no stand on whether, or how often, things are this bad with regard to the moral/political. I find it hard enough to think about “what happened” in a situation where two chipmunks chased each other. (Though if our natural ways of describing action can quickly lead to puzzles when we talk about a situation in which each of two animals acts in a way that targets the other, then we may need to develop some other ways of talking about morally complex situations.)

In any case, when considering (9) and (10), we need to be realistic about human minds.

\begin{align}
(9) & \text{Alvin chased Theodore gleefully and athletically but not skillfully.} \\
(10) & \text{Theodore chased Alvin gleefully and unathletically but skillfully.}
\end{align}

To represent is to represent in a format; and formats can have “side-effects” that are not negligible. Some of our “intuitions” may reflect multiple factors that include deep framing effects, as in Schelling’s case, and not “mere” laziness effects (as in the bat and ball example). There is no

---

7 For any income \( i \): Base(i) – \( k \) [Deduction(i)] = Base* (i) + (2 – \( k \ )) [Surcharge(i)]; and Base* (i) = Base(i) – 2 [Deduction(i)]; hence, – \( k \ ) [\text{Deduction}(i)] = – 2 [\text{Deduction}(i)] + (2 – \( k \ )) [\text{Surcharge}(i)] ; so (2 – \( k \ )) [\text{Deduction}(i)] = (2 – \( k \ )) [\text{Surcharge}(i)]; so Deduction(i) = Surcharge(i).
guarantee that such intuitions can be coherently described as having stable propositional contents; cp. Kripke (1979). This matters for purposes of evaluating the hypothesis that sentences like (9) and (10) have truth conditions and logical forms along the lines of (9a) and (10a).

\[(9a) \exists x[\text{chased}(x, \text{AL}, \text{THEO}) \& \text{gleeful}(x) \& \text{athletic}(x) \& \neg \text{skillful}(x)]\]

\[(10a) \exists x[\text{chased}(x, \text{THEO}, \text{AL}) \& \text{gleeless}(x) \& \text{unathletic}(x) \& \text{skillful}(x)]\]

The intuitions that motivate appeals to these eventish logical forms may reflect details of linguistic description, as opposed to description-neutral (truth-conditional) content.

By uttering (9) or (10), a speaker can correctly report some of what happened when one chipmunk chased another. But no one event can satisfy all eight of the conjuncts in (9a) and (10a). So one might conclude that if (9) and (10) have the indicated logical forms, then the things that happened included two chases involving Alvin and Theodore. This sets us up for a familiar kind of debate that nobody wins. We can, however, reject the conditional.

Instead of viewing (9a) and (10a) as specifications truth conditions that (9) and (10) allegedly have, we can view the invented sentences as initial models of thoughts that might be assembled—given the right ingredients—by executing the instructions that (9) and (10) provide. Thoughts actually constructed via (9) and (10), on particular occasions, may or may not rise to the level of being true or false. Such thoughts may also be less sophisticated than (9a) and (10a), which employ Tarskian variables and conjoiners. We may use (9) and (10) to form thoughts, perhaps isomorphic to (9a) and (10a), that are not “made true” by any event that we are thinking about in two (perhaps slightly distorted) ways. We may represent and report what happened in existential form, though in framed ways that preclude language-independent events from being values of the relevant variables. Defending this positive proposal requires the details in chapters five and six. But in any case, one can adopt “eventish” specifications of what (3) and (4) mean without adopting the hypothesis that (3) and (4) have the truth conditions specified by the formal sentences used to specify the linguistic meanings.

3. The Uncomfortable Event Position

One might prefer to jettison event analyses. But we still need to explain why the inference from (2) to (1) is compelling. Davidson’s original argument can also be extended.

(2) Al chased Theo gleefully.  
(1) Al chased Theo.

3.1 Compositionally Indispensable

Shifting examples, suppose that Miss Scarlet stabbed Colonel Mustard twice. One stab was done with a grey dagger; it was a bit clumsy, resulting in a superficial wound on Mustard’s left side. The other one, done with a red dagger, was a proficient and fatal stab on his right side. Given this context, Davidsonian analysis yields a network of correct predictions concerning (28-35).

(28) Scarlet stabbed Mustard clumsily with a grey dagger.  
(29) Scarlet stabbed Mustard proficiently with a red dagger.  
(30) Scarlet stabbed Mustard clumsily.  
(31) Scarlet stabbed Mustard with a grey dagger.  
(32) Scarlet stabbed Mustard proficiently.  
(33) Scarlet stabbed Mustard with a red dagger.  
(34) Scarlet stabbed Mustard clumsily with a red dagger.  
(35) Scarlet stabbed Mustard proficiently with a grey dagger.

The conjunction of (28) and (29) implies each of (30-33), but not (34) or (35).  

Let S(x) abbreviate \text{stabbed}(x, \text{SCARLET, MUSTARD}). Let C(x)/P(x)/G(x)/R(x) abbreviate \text{clumsily}(x)/\text{proficiently}(x)/\text{with-a-grey-dagger}(x)/\text{with-a-red-dagger}(x).
Call the two stabbings Gauche and Droite, each of which (is the value assigned to the variable by an assignment that) satisfies S(x). Gauche satisfies C(x) and G(x), but not P(x) or R(x); Droite satisfies P(x) and R(x), but not C(x) or G(x). So (28a) and (29a) are true.

\[(28a) \exists x[S(x) \& C(x) \& G(x)]\]
\[(29a) \exists x[S(x) \& P(x) \& R(x)]\]

Hence, \(\exists x[S(x) \& C(x)]; \exists x[S(x) \& G(x)]; \exists x[S(x) \& P(x)];\) and \(\exists x[S(x) \& R(x)].\) But the logical forms (34a) and (35a) are false.

\[(34a) \exists x[S(x) \& C(x) \& R(x)]\]
\[(35a) \exists x[S(x) \& P(x) \& G(x)]\]

In short, (28-35) exhibit a pattern of implications and nonimplications that is expected, given the conjunction reduction analysis illustrated with (2a) and (1a).  

\[(2a) \exists x[CHASED(x, AL, THEO) \& GLEEFUL(x)]\]
\[(1a) \exists x[CHASED(x, AL, THEO)]\]

I don’t know of any other explanation for this pattern. We can introduce more dyadic notions: STABBED-CLUMSILY(X, X’); STABBED-CLUMSILY-WITH-A-GREY-DAGGER(X, X’); STABBED-CLUMSILY-WITH-A-RED-DAGGER(X, X’); STABBED-WITH-A-GREY-DAGGER(X, X’); etc. But analyzing (28-35) in terms of such notions doesn’t explain why every satisfier of STABBED-CLUMSILY-WITH-A-GREY-DAGGER(X, X’) satisfies both STABBED-CLUMSILY(X, X’) and STABBED-WITH-A-GREY-DAGGER(X, X’)—though not conversely. It seems that the meaning of ‘stab’ somehow indicates a conceptual slot corresponding to stabs, as indicated in (28a) and (29a), whose conjunction implies neither (34a) nor (35a).

Note, however, that satisfiers of S(x) have to be individuated more finely than ordered triples of the form \(<t, Scarlet, Mustard>\); where \(t\) is some moment (or interval) of time. The two stabs of Mustard by Scarlet may have been simultaneous, each occurring at dawn. By itself, this isn’t yet a problem. One can plausibly posit simultaneous events that have the same participants. Famously, a sphere might be heating up as it spins; see Kim (1976), Davidson (1969). And while Scarlet participated in both Gauche and Droite, it might be that she moved her left hand in the former event as she moved her right hand in the latter. Similarly, while Mustard participated in both events, his right side may have been affected in Gauche as his left side was affected in Droite; the daggers did not simultaneously make a single wound. So this isn’t yet an objection to Davidonian analysis. On the contrary, if all the examples were like this, I would be more favorably disposed towards truth-conditional semantics. But it’s not enough to individuate events more finely than n-tuples of times and participants.

Imagine two rocks colliding, exactly once, with (36) and (37) being correct reports.

\[(36) \text{ The grey rock struck the red rock.}\]
\[(37) \text{ The red rock struck the grey rock.}\]

Call the rocks Grey and Red. It is tempting to say that a single event—the collision of Grey and Red, a.k.a. the collision of Red and Grey—satisfies both STRUCK(X, GREY, RED) and the distinct open sentence STRUCK(X, RED, GREY). But consider (38) and (39).

---

8 Taylor (1985) discusses such examples, citing Christopher Arnold who cited Gareth Evans. When we’re not worried about whether sentences can be true, we might say that (28-33) can be true while (34-35) are false. When we are worried, we can say that there are contexts in which speakers can tell truth with (28-33) but not (34-35). Of course, the conjunction-reduction explanation does not require the existence of Gauche and Droite. But in this case, Tarskian analogs of (18-23) may well be true. Gauche and Droite are, plausibly, distinct events.

9 But we cannot identify both stabs with \(<\text{dawn, Scarlet, Mustard}>\). As Evans stressed, events need to happen at times and have participants, but without being n-tuples of times and participants.
The grey rock struck the red rock forcefully from the west.

The red rock struck the grey rock forcefully from the east.

Prima facie, the collision does not satisfy FROM(X, THEWEST) or FROM(X, THEEAST).

In response, one might say that STRUCK(X, X', X'') abbreviates STRUCK(X) & ∃E[X = <E, X', X'']].

On this view, one can say both that STRUCK(X, X', X'') is satisfied by assignment A if A(e) was an event of A(x) striking A(x')—or more briefly, that the open sentence is true of <e, α, β> iff e was an event of α striking β—and that an event of α striking β can be one of β striking α. If the logical forms of (36) and (37) are (36a) and (37a),

(36a) \( ∃x \{ \text{struck}(x) \& \exists e[x = <e, \text{grey}, \text{red}>] \} \)

(37a) \( ∃x \{ \text{struck}(x) \& \exists e[x = <e, \text{red}, \text{grey}>] \} \)

then the collision can be the one relevant value of e for both sentences. But there are two relevant values of x: <the collision, Grey, Red> and <the collision, Red, Grey>. In this sense, (36a) and (37a) have distinct “truth makers;” the one satisfier of \( \text{struck}(x) \& \exists e[x = <e, \text{grey}, \text{red}>] \) differs from the one satisfier of \( \text{struck}(x) \& \exists e[x = <e, \text{red}, \text{grey}>] \).

One can go on to say that adverbial modifiers are like verbs in being true of ordered n-tuples that include participants. Perhaps FORCEFULLY(X) is satisfied by both triples, while FROM(X, THEWEST) is only satisfied by the first, and FROM(X, THEEAST) is only satisfied by the second. This accommodates (38-39). But the proliferation of truth makers suggests that linguistic descriptions do matter, after all, with regard to what the conjuncts of logical forms are true of. Moreover, the issues illustrated with (38-39) are not confined to overtly perspectival predicates.

Davidson took the values of his event variables to be both things that can be described in many conjunctive ways, at least in an ideal language, and things that many human language predicates are true of. The question is whether our naturally acquired words conform to this conception of them. This invites attention to episodes that can be described in grammatically distinct but symmetric ways, as in (36-37), or in a correspondingly neutral way as in (40-41).

(40) A rock struck a rock forcefully.

(41) There was a collision.

One wants to say that a certain language-independent event, the collision of the rocks, can be the one thing that \( ∃x ∃x' \{ \text{struck}(x, x', x'') \& \text{rock}(x') \& \text{rock}(x'') \& \text{forcefully}(x) \} \) is true of. But was the striking of Red (by Grey) exactly as forceful(ly) as the collision? More generally, is every English predicate true of the collision if and only if it is true of either striking of rock by rock? Answering affirmatively invites a parade of apparent counterexamples. Answering negatively undermines the attractions of Davidsonian analyses if we presuppose thesis (L).

(L) if \( S \) is a sentence of a Human Language, and \( S* \) is the logical form of \( S \),

then \( S \) has the truth condition that \( S* \) specifies

So perhaps we should reject (L) and view (36-41) as illustrations of framing, as opposed to existential closures of various event predicates that are true of a single event.

3.2 Ontologically Valueless

Examples like (42) raise similar issues, though with further complications.

(42) Mister Green married Miss Scarlet enthusiastically, but

Miss Scarlet married Mister Green unenthusiastically.

---

One might say that MARRIED(X, GREEN, SCARLET) and MARRIED(X, SCARLET, GREEN) are true of different events, done by Green and Scarlet respectively. But if one of these events ended before the other—say because each event of marrying is identified with a certain speech act, and Green spoke first—there is an obvious difficulty. For even if (43) is correct, (44) is not.

(43) Green spoke before Scarlet did.
(44) Green married Scarlet before she married him.

I think that (43a) and (44a) are plausible Davidsonian logical forms.

(43a) \( \exists x \) [SPOKE(x, GREEN) & \( \exists x' \) [BEFORE(x, x') & SPOKE(x', SCARLET)]]
(44a) \( \exists x \) [MARRIED(x, GREEN, SCARLET) & \( \exists x' \) [BEFORE(x, x') & MARRIED(x', SCARLET, GREEN)]]

But if (44a) is false, then (42) presents the same kind of difficulty as (3) and (4).

(9) Al chased Theo gleefully and athletically but not skillfully.
(10) Theo chased Al gleefully and unathletically but skillfully.

Either MARRIED(X, GREEN, SCARLET) and MARRIED(X', SCARLET, GREEN) are satisfied by the same event, making it hard to see how (42) could be true, or Green and Scarlet participated in two simultaneous events of marrying.

Note that Scarlet and Green become married only if they both participated actively in the ceremony. Likewise, both chipmunks had to be active in any event of one chasing the other; you can’t chase a chipmunk that stays still. But (1) represents Al as more agentive than Theo,

(1) Al chased Theo.

while (3) represents Theodore as more agentive.

(3) Theo chased Al.

So CHASED(X, AL, THEO) might be elaborated as in (45), and likewise for CHASED(X, THEO, AL).

(45) AGENT(X, AL) & PAST(X) & CHASE(X) & PATIENT(X, THEO)]

Then even without adverbs, it seems that no one event can be both Al’s chase of Theo and Theo’s chase of Al. Al is the agent of the former, and Theo is the Patient of the latter. Similarly, it seems that no event of Scarlet chasing Green could be an event of Green fleeing Scarlet.

In response, one might claim that an event can have more than one agent, and that \( \varepsilon \) satisfies AGENT(X, AL) if Al is one of the one or more agents of \( \varepsilon \). But then it seems that (46)

(46) Alvin lifted the piano, and then he played the trio.

would be true if Al was one of three musicians who together lifted the piano and played the trio.\(^{11}\) That’s wrong. So plausibly, \( \varepsilon \) satisfies AGENT(X, AL) only if Al is the agent of \( \varepsilon \). So it’s hard to see how Al’s chase of Theo could be Theo’s chase of Al. Reports of perceptions, as in (47-49), raise a similar issue while providing further evidence for eventish logical forms.

(47) Peacock heard Mustard yell.
(48) Peacock heard a yell.
(49) Mustard yelled, and Peacock heard him.

Such reports remind us that a verb like ‘heard’ can combine with an untensed clausal complement, like ‘Mustard yell’, to form a phrase.\(^{12}\) It is hard to accommodate this point if the logical form of ‘heard’ is simply HEARD(X, Y). But if the logical form of (49) is (49a),

\(^{11}\) See Schein (2002). Perhaps Alvin did some piano-lifting. But (46) implies that some event was a lifting of the piano by Alvin. Moreover, ‘the piano’ can be replaced with ‘five pianos at once’.

\(^{12}\) The phrase ‘heard that Mustard yelled’ is different again. If (47) is true, and Mustard was the tallest officer, then Peacock heard the tallest officer yell. But if she heard that Mustard yelled—say, because Plum passed on the rumor—she need not have heard that the tallest officer yelled.
Suppose that Scarlet shot Green exactly once, shortly after marrying him. In doing so, Scarlet will have acted in a way that can be reported in many ways, say, with (55) and (56).

(55) Scarlet shot Green with her revolver.
(56) Scarlet pulled the trigger with her ring finger.

4. Trying to Restore Comfort
I grant that some apparent puzzles for event analyses can be dealt with via two strategies: posit distinct but related events (§4.1); or posit some relativization to description (§4.2). But in my view, neither strategy plausibly extends to the hard cases.

4.1 Distinguish but Relate
Suppose that Scarlet shot Green exactly once, shortly after marrying him. In doing so, Scarlet will have acted in a way that can be reported in many ways, say, with (55) and (56).

(55) Scarlet shot Green with her revolver.
(56) Scarlet pulled the trigger with her ring finger.
But let’s suppose that (57) and (58) are not correct ways of reporting what happened.

(57) Scarlet shot Green with her ring finger.

(58) Scarlet pulled the trigger with her revolver.

If (55a) and (56a) are true, but (57a) and (58a) are not,

Then SHOT(X, SCARLET, GREEN) and PULLED(X, SCARLET, THETRIGGER) are satisfied by distinct events; see, e.g., Pietroski (1998). That’s not surprising. The event sortals differ. Green differs from the trigger. And the trigger was affected first, with the subsequent result that Green was affected; cp. Feinberg (1965) on the “accordion effect.” So one might say that the event of Scarlet pulling the trigger was part of the event (or process) of her shooting Green, which may have been part of her killing Green; see Thalberg (1972), Thomson (1971, 1977). The trigger-pulling may itself have included a certain motion of Scarlet’s ring finger, and perhaps a mental cause of that bodily motion. But here, it’s enough to note that Green wasn’t shot until the bullet entered him. The trigger was pulled a moment before. So plausibly, SHOT(X, SCARLET, GREEN) and PULLED(X, SCARLET, THETRIGGER) are true of distinct but related events.

I think it’s less plausible that one of Scarlet’s actions is the truth maker for both (55a) and (56a). Davidson (1967a) tried this “identificationist” strategy, according to which the apparent falsity of (57a) and (58a) must be explained away, in response to examples like (59) and (60).

(59) Scarlet shot Green at dawn, and he died (later that day) at noon.

(60) Scarlet killed Green at dawn, and he died (later that day) at noon.

Intuitively, (59) can be a correct report of what happened, while (50) is not. But Davidson held that if (59a) is true, so is (60a): (60) just sounds wrong because Green wasn’t dead at dawn.

I think it’s less plausible that one of Scarlet’s actions is the truth maker for both (55a) and (56a). Davidson (1967a) tried this “identificationist” strategy, according to which the apparent falsity of (57a) and (58a) must be explained away, in response to examples like (59) and (60).

(59) Scarlet shot Green at dawn, and he died (later that day) at noon.

(60) Scarlet killed Green at dawn, and he died (later that day) at noon.

Intuitively, (59) can be a correct report of what happened, while (50) is not. But Davidson held that if (59a) is true, so is (60a): (60) just sounds wrong because Green wasn’t dead at dawn.

(59a) ∃x[SHOT(X, SCARLET, GREEN) & AT(X, DAWN) & ∃x‘[DIED(X’, HE) & AT(X’, NOON)]]

(60a) ∃x[KILLED(X, SCARLET, GREEN) & AT(X, DAWN) & ∃x‘[DIED(X’, HE) & AT(X’, NOON)]]

The idea was that while STABBED(X, SCARLET, GREEN) and KILLED(X, SCARLET, GREEN) are both satisfied by Scarlet’s action at dawn, her action is not correctly described as a killing until Green dies, much as an award-winning performance is not correctly described as such before the award is given. According to Davidson, an action occurs wholly where the actor is. If the action was a killing, it follows that someone died; though it doesn’t follow the death was part of the action. I agree. But one can grant this point, taking actions to be causal contributions of agents, and conclude that KILLED(X, SCARLET, MUSTARD) is not satisfied by any action. For one might think that utterances of (50) are false, and not merely odd, given that (61) sounds fine.

(61) Scarlet shot Green fatally at dawn, and he died at noon.

A stabbing, like an illness, can be described as fatal before the relevant death. So it seems that the killing differs from the fatal stabbing after all: the killing, unlike Scarlet’s action, is not

---

13 See Wilson (1989), Ginet (1990), Pietroski (2000). Perhaps (59) and (60) can be used to talk about the same “root action,” whatever it turns out to be. But speakers can use expressions to talk about things that the expressions are not true of; see, e.g., Donnellan (1966). One can posit forms like ∃α∃x[R(α, X) & SHOT(X, SCARLET, GREEN)]; where R(α, X) is true of <α, β> if and only if α is the action “at the root of” β. But even if this formal claim is true if and only if Scarlet shot Green, this may reflect the metaphysics of shooting, not the meaning of ‘shoot’.
over until Green dies. So positing *distinct but related* events is plausible for some cases, especially where it is independently plausible that the relevant event predicates are satisfied by processes that exhibit part-whole relations. But this strategy seems especially unsuited to examples like (9) and (10).

(9) Alvin chased Theodore gleefully and athletically but not skillfully.
(10) Theodore chased Alvin gleefully and unathletically but skillfully.

If there are two chasings, they occupy the same region of spacetime and have the same participants. Such events, satisfying the same sortal term, would be very intimately related.

Perhaps one can distinguish a statue from the lump of material that constitutes it, even if their careers (as existents) start and finish together. For perhaps statues and lumps of material are just things of different sorts, and this underpins certain modal differences. But it begs the questions at hand to insist that Alvin’s chase of Theodore differs modally from Theodore’s chase of Alvin. One can try to argue that chases of Theodore and chases of Alvin are events of different sorts. But then one needs some account of event sorts. Gleeful chases of Theodore and athletic chases of Theodore had better not be events of different sorts, in any way that precludes (token) identity claims, on pain of spoiling Davidsonian accounts of implications and nonimplications. One can say that “direct objects matter” for these purposes. But then it’s hard to see the difference between positing two chases, and positing one chase along with the claim that a certain difference in linguistic description is semantically relevant.

In this respect, examples like (52) and (53) are again instructive.

(52) Simon played the song on his tuba.
(52a) ∃x[PLAYED(x, Simon, TheSong) & ON(x, HisTuba)]
(53) Simon played his tuba.
(53a) ∃x[PLAYED(x, Simon, HisTuba)]

Suppose that Simon played the song in question exactly once. Let π be the event of him playing that song, so that π is the truth maker for (52a) and plausibly a truth maker for (53a). Even if (52a) does not logically imply (53a), the latter is true if the former is true. Of course, (53a) would be true if Simon only played part of the song on his tuba. But let π’ be the event of tuba playing that has the same spatiotemporal properties as π. Then π’ is one among many truth makers for (53a). And intuitively, π’ is the same event as π. If π was a performance of “My Way” that started at noon and ended three minutes later, then π’ can also be described this way. Moreover, π seems to have all the same causes and effects as π’: if some vibration of air was caused by π, then it was caused by π’; if π was caused by a desire to impress the audience, then so was π’.

It might be noted that Simon played the song by playing his tuba, not vice versa. But this asymmetry, which may well reflect order of intentions, does not show that π and π’ are distinct events; cp. Anscombe (1957), Thomson (1977), Hornsby (1980). So I wouldn’t want to rely on ‘by’, which presents complications for everyone, in arguments against Davidsonians.

It does, however, beg the questions at hand to insist that the song-playing π differs modally from the tuba-playing π’. And insisting that “direct objects matter” will lead to grief. For suppose that Simon flipped a certain switch and thereby played: a record, a song, a Beatles tune, a cover of a Beatles Tune, his favorite track, a recording of a song, a hit record, a top ten hit. (If Simon is a deejay, he may have played all those things on the radio.) It seems perverse to say that Simon is the agent of so many distinct but related events, occupying the same region of spacetime, instead of saying that these are many ways of reporting some of what happened when he flipped the switch. Reporters can choose from boundlessly many different descriptions of certain effects of making a certain grooved piece of vinyl spin.
I can’t prove that the strategy of positing distinct but related events is wrong for these cases. But this strategy does not merely suffer from metaphysical profligacy. It threatens the very idea that variables have values that can satisfy boundlessly many predicates, many of which are logically independent. And if this idea is abandoned, it’s hard to see point of saying that human language predicates are often satisfied by events. This is not to deny the attractions of positing distinct but related events. Many reports of “what happened” in a given region of spacetime may have the logical form \( \exists x [\Phi(x)] \), for some complex predicate ‘\( \Phi \)’. But it may not be possible to identify any event as both the relevant “thing that happened” and the common truth maker for (the logical forms of) all the reports. This can make it tempting to posit many events—at the limit, one per report—until one thinks about the implications.

If we could be sure that our judgments in these cases coherently reflect the things that happened, however many there were, then our judgments concerning the truth of sentential utterances might favor the strategy of positing surprisingly many events. But in using human languages to report and think about what happened, we may report and think only in frame-dependent ways. And there is no guarantee that our various ways of framing what happened can be coherently viewed as representations of description-neutral events. Indeed, we have independent evidence of framing effects that run deep in human cognition, but no independent evidence for co-located playings/chasings/etc. So absent good reasons for thinking that our judgments in these cases reflect the existence of distinct but intimately related events, positing such events seems less motivated than revising the relevant assumptions about how meaning, truth, and logical form are related. But there remains the question of whether such revision is better motivated than less metaphysically profligate ways of responding to the puzzle cases.

4.2 Identify but Relativize
Consider again the event \( \pi \) of Simon playing the song on his tuba, and the corresponding event \( \pi' \) of Simon playing his tuba. As noted above, simply identifying \( \pi \) with \( \pi' \) won’t do.

Suppose that \( \pi \) satisfies \( \text{played}(x, \text{Simon, HisTuba}) \) and \( \text{played}(x, \text{Simon, TheSong}) \) and \( \text{on}(x, \text{HisTuba}) \). Then \( \pi \) is a truth maker for (54a). In which case, (54) is true

(54) Simon played his tuba on his tuba.

(54a) \( \exists x [\text{played}(x, \text{Simon, HisTuba}) \& \text{on}(x, \text{HisTuba})] \)

If this human language sentence has the truth condition specified by (54a). But consider (62-63).

(62) Simon played the song in three minutes.

(62a) \( \exists x [\text{played}(x, \text{Simon, TheSong}) \& \text{in-three-minutes}(x)] \)

(63) Simon played his tuba for three minutes.

(63a) \( \exists x [\text{played}(x, \text{Simon, HisTuba}) \& \text{for-three-minutes}(x)] \)

If (62a) and (63a) are true, then \( \pi \) satisfies \( \text{in-three-minutes}(x) \) and \( \text{for-three-minutes}(x) \). In which case, there is presumably some three-minute interval such that: \( \pi \) occurred within that interval; and \( \pi \) lasted for that interval. But then (64a) is equally true. Yet the deviant (64) is not.

(64) *Simon played his tuba in three minutes.

(64a) \( \exists x [\text{played}(x, \text{Simon, HisTuba}) \& \text{in}(x, \text{three minutes})] \)

This raises the question of why (64) cannot be understood as a way of expressing (64a). By way of comparison, recall that while (65) is defective, it has the meaning of (65a).

(65) *The child seems sleeping.

(65a) The child seems to be sleeping.

(65b) #The child seems sleepy.
So why isn’t (64) understood as a (perhaps odd) way of saying what one says with (63)?

One can describe the unacceptability of (64) in terms of a formal property exhibited by ‘in’ but not ‘for’. Indeed, the basic facts are well known. Roughly speaking, a “telic” modifier like ‘in three minutes’ is licenced only if the modified phrase specifies an “endpoint” condition that determines when an event of the relevant sort has terminated. Compare (66) with (67).

(66) Simon ran to/through/past the park in three minutes.

(67) *Simon ran towards/in/near the park in three minutes.

An event of running to the park ends when the runner reaches the park, while ‘towards the park’ has no corresponding implication, even though every run ends somewhere. Note that ‘ran around the park in an hour’ implies a loop, but ‘ran around the park for an hour’ does not. Likewise, an event of playing the song ends (modulo vagueness) when the final notes are sounded. But ‘played his tuba’ has no corresponding implication, even though no tuba playing goes on forever.

That, however, is no defense of the idea that ‘played his tuba’ and ‘in three minutes’ are satisfied by language-independent events. On the contrary, it suggests that meanings have to do with how events are represented. The triviality of (68) suggests that in some sense,

(68) If Simon played the song on his tuba in three minutes,

then Simon played his tuba for three minutes.

speakers recognize that any event of playing a song on a tuba in three minutes is an event of playing that tuba for three minutes. Yet replacing ‘for’ with ‘in’—the very preposition used in the antecedent of (68)—makes a mess of the consequent.

(69) If Simon played the song on his tuba in three minutes,

then Simon played his tuba in* three minutes.

This can be explained if ‘in three minutes’ is unlike \( \text{IN}(X, \text{THREE MINUTES}) \), in that the human language phrase is a grammatically coded instruction for how to build a certain kind of telic concept, while ‘played his tuba’ is a grammatically coded instruction for how to build a certain kind of atelic concept. But if ‘in three minutes’ is like \( \text{IN}(X, \text{THREE MINUTES}) \) in being satisfied by \( \pi \) if and only if \( \pi \) took place in three minutes, then why can’t (68) and (69) be so understood?

The point here is not limited to ‘in’ and ‘for’. Suppose that Simon used all the available polish to clean some tarnished brass, getting the job done in an hour. One might describe Simon’s contribution in any of the four ways indicated with (70-71).

(70) Simon put the polish on brass for/in an hour.

(71) Simon polished the brass for/in an hour.

We have the usual motivations for identifying the event of putting the polish on brass with the event polishing the brass. But if the direct object is a bare noun, ‘in an hour’ is no longer a viable way of extending the predicate.

(72) Simon put polish on (the) brass for/* in an hour.

(73) Simon polished brass for/* in an hour.

---

14 One has the feeling that an event of playing a tuba is somehow divisible into shorter episodes of that same kind, in a way that no event of playing a song is—at least until one thinks about the possibility of playing a one-note song. But even if the distinction is hard to specify, and more mental than metaphysical, one might think that it bears on why Simon’s tuba playing is not said to be in two minutes; perhaps the tuba-playing satisfies \( \text{IN}(X, \text{TWO MINUTES}) \); but it is pragmatically odd to say so, since ‘in two minutes’ suggests events that are not represented as indefinitely divisible. The relevant implicature would be uncancelable and unaffected by conditionalization. But even if this explains why it is weird to use (64), the question is why (64) cannot be understood as having the meaning of (63).

15 See especially Tenny (1994) and further references there.
While every brass-polishing eventually ends, thinking about an event ε as one of putting polish on the brass is not to think about ε as an event that includes a culminating point at which the brass is polished; and likewise for thoughts of polishing brass. But once the polish is on brass, or the brass is polished, the job is done. Sentence (70) has us thinking about the “job” as using up the polish, even if Simon had no such intention, while (71) has us thinking about clean brass as the telos. If this distinction correlates with the form of the direct object—bare noun or determiner phrase, even when the same event/performance/work/contribution is variously described—one begins to suspect that Davidsonian “event variables” are best viewed as part of a theory of how linguistic expressions are related to human concepts, without regard to the events that might be thought about and/or stipulated to be values of variables in some invented ideal language.16

Given these considerations, one might supplement an identification strategy with some appeal to description-relative predicates. In particular, one might draw attention to an example that Davidson (1967a) discussed. Suppose that Jones swam across the English Channel. Presumably, an event can be quick for a channel-swimming yet slow for a channel-crossing, much as an insect can be large for an ant yet small for an animal. So it would be rash to conclude that the swimming was an event distinct from the crossing, just because both (74) and (75)

(74) Jones swam the channel quickly.
(75) Jones crossed the channel slowly.

can be used to report what Jones did. Formalism like (74a) and (75a) can be viewed as shorthand for more articulated logical forms of the sort indicated in (74b) and (75b).

(74a) ∃x[swam(x, Jones, TheChannel) & quick(x)]
(74b) ∃x[agent(x, Jones) & swam(x, TheChannel) & quick-for(x, _)]
(75a) ∃x[crossed(x, Jones, TheChannel) & slow(x)]
(75b) ∃x[agent(x, Jones) & crossed(x, TheChannel) & slow-for(x, _)]
The (b)-forms encode the idea that ‘quickly’ and ‘slowly’ are relativized to ‘crossed the channel’ and ‘swam the channel’, allowing for comparisons to other events of the same sort, perhaps done by other agents. And I have no objection to separating out ‘agent(x, Jones)’.

This suggests an obvious extension to examples like (76) and (77); see §2.4 of chapter five.

(76) Simon played his tuba well.
(77) Simon played the song well.

16 One can introduce a predicate of concepts, culminater(_), that applies to play-a-song and put-the-polish-on-brass but not play-a-tuba or put-polish-on-the-brass. But such a predicate will not apply to Simon’s performance. One can introduce a “secondary quality” predicate, sq-culminater(_), that applies to events that can be thought about with a concept to which culminater( _) applies. But then sq-culminater(_) applies to the event of Simon playing his tuba, since that event can be thought of as the playing of a song. Note that (70-73) also tell against the idea that in ‘played his tuba’ and ‘played the song on his tuba’, ‘his tuba’ is semantically related to the same thematic concept (while ‘the song’ is semantically related to another thematic concept.) On such a view, either ‘play’ lets the thematic concept associated with its direct object vary, or ‘his tuba’ is not the direct object in ‘played his tuba’. But the first option would violate a well confirmed “uniformity” constraint, discussed in chapter five, on how thematic concepts are grammatically expressed; see Baker (199x) and references there. The second option seems wrong for many reasons, including the fact that ‘played the tubas in an hour’ is much better than ‘played (on the) tubas in an hour’, as if ‘the tubas’ is the direct object and telos-specifier in the former but not the latter. (Pluralizing helps only if the direct object is a determiner phrase.) If in-an-hour can( _) only extend concepts that represent processes as telic and extended, the deviance of ‘played/pounded the tuba in an hour’ is expected, even with a determiner phrase as direct object. But ‘pounded/played the tubas in an hour’ is presumably better because pluralizing provides a way of thinking about the process as extended.
Taking ‘well’ to be the adverbial version of ‘good’, one might say that an event can be good for a tuba-playing without being good for a song-playing. But if only because ‘big’, ‘slow’, and ‘good’ can appear in constructions like (78),

\( \text{(78) Of those two ants, Adam is the big/slow/good one, but other ants are even bigger/slower/better.} \)

it isn’t clear how to plausibly extend the “identify but relativize” strategy to other modifiers. Put another way, it’s not implausible that ‘big’, ‘slow’, and ‘good’ are understood in terms of the relational concepts that we often express with comparative morphemes. To a first, though only first approximation: to be a big/slow/good \( \Phi \) is to be a \( \Phi \) that is bigger/slower/better than most.\(^{17}\) And if \( x \) is big/slow/good, then for some \( \Phi \), \( x \) is a big/slow/good \( \Phi \); cp. Thomson (2008). But prima facie, this pattern does not extend to all modifiers, lexical or phrasal. Consider (79).

(79) Simon played the song in the style of Sinatra.

I don’t think that ‘in the style of Sinatra’ is understood, on its own, in terms of what it is to be in the style of Sinatra for a \( \Phi \). But one can hypothesize that the relativization is introduced compositionally, as a reflex of modifying a phrase with a grammatical adjunct. Perhaps all phrases of the form [[VERB OBJECT] MODIFIER] have the logical form indicated in (80),

\( \text{(80) VERB}(x, \text{OBJECT}) \& \text{MODIFIER-FOR}(x, \_i) \)

whose instances include ‘Played,(e, the song) \& In-The-Style-of-Sinatra-For(e, \_i)’. But note that this changes—and arguably trivializes—the Davidsonian account of predicate modification: it’s not that modifiers apply to the very events that unmodified verb phrases apply to; rather, modifiers express relations that can be exhibited by verb phrases and whatever they apply to. Moreover, for many modifiers as in (81), any relativization has to be truth-conditionally inert.

(81) Yesterday, Simon played his tuba while Phosphorus was still visible. Recall the similar point concerning ‘struck forcefully’ vs. ‘struck from the east’.

It’s also hard to know what the “identify but relativize” strategy predicts, once we get beyond examples like (74-77). If a theory predicts that (79) is true only if some event of Simon playing the song satisfies a certain metalanguage predicate \( P \), then one wants to know if Simon’s song-playing satisfies \( P \). Now it’s already a little unclear what it is for an event to be in the style of Sinatra. But as a competent speaker of English who knows what a personal style is, I could—perhaps after reviewing some of Sinatra’s performances—make judgments about whether certain performance by others were Sinatrasque. I would, however, be confused by the following instruction: don’t say whether what Simon did on stage (with his tuba) was done in the style of Sinatra; rather, say whether Simon’s playing of the song was Sinatrasque for an event of that sort, and whether his tuba-playing was Sinatrasque for an event of that sort. So even if the logical form of (79) is more complex than (79a), (79b) seems complex in a wrong way.

\( \text{(79a) } \exists x[\text{PLAYED}(x, \text{SIMON}, \text{HIS TUBA}) \& \text{IN}(x, \text{THE STYLE OF SINATRA})] \)

\( \text{(79b) } \exists x[\text{AGENT}(x, \text{SIMON}) \& \text{PLAYED}(x, \text{THE SONG}) \& \text{IN-THE-STYLE-OF-SINATRA-FOR}(x, \_i)] \)

Again, my point is not that relativizing to descriptions is always illegitimate. A musical performance can trigger the spontaneous judgment that the instrument was played well, while the

\(^{17}\) Cp. Ross (1967), Klein (198x). See also Kennedy (199x, 200x) and Schwartzschild (200x) for proposals, in terms of scales, and further references. See Pietroski (2006b) for discussion of the formal relation to Frege’s notion of an ancestral, Boolos’ (1998) plural interpretation of second-order logic, and a corresponding “Conjunctivist” analysis of ‘big ant’.
music was not. We know that a four-minute mile is quick in some contexts but not others. The quick/slow contrast can also be deployed to modify boundlessly many intuitively diverse predicates—e.g., ‘crossed the Channel’, ‘proved a theorem’, and ‘removed an appendix’. Correlatively, we tacitly know what it is to be quick for a Channel-swimming yet slow for a Channel-crossing: it is (roughly) to be an event of swimming across the Channel in significantly less than the average time required for such a swim, though significantly more than the average time required to get across the Channel in a more standard way. So plausibly, speakers know that this is what it is for a value of an event variable to be one in which the Channel is swum quickly but crossed slowly. Many other dimensions of variation can be relativized to predicates in this way. But prima facie, boundlessly many modifiers do not fit this model.

I think we know—or at least can come to know—what it is for a vocal performance, or an acting performance, or perhaps a whole life to be Sinatraesque. And we may know what it is for a performance to be Sinatraesque for a vocal (as opposed to acting) performance. But it doesn’t follow that we know what it is for a singing of “Happy Birthday” to be Sinatraesque for a singing of that song, as opposed to being Sinatraesque simpliciter. Given Marilyn Monroe’s famous singing of that song for John Kennedy, we know how to construe (82).

(82) At the party, someone sang “Happy Birthday” in the style of Monroe. But I don’t think that (83) implies a distinctly Sinatraesque version of that particular song.

(83) At the party, someone sang “Happy Birthday” in the style of Sinatra.

Likewise, we may have a dim sense of what it would be to prove a theorem (remove an appendix, play a tuba) Sinatraesquely—perhaps it involves holding a cigarette, wearing a Fedora at just the right angle, and making the task at hand seem easy—without having any independent sense of what it would be for an action to be Sinatraesque for a theorem-proving. In short, I think that apart from some important special cases, the relativization component of the “identify but relativize” strategy is otiose. It does provide formalism that insulates Davidsonian event analyses from certain objections. But that is not a virtue. Let me change the example to stress this point.

Suppose that Simon played the song in the way his Uncle Jim taught him: in a certain key, with a certain lilt, and introducing a few unexpected notes during the bridge. Suppose that Simon also played his tuba in the way his Aunt Joan taught him: with a certain embouchure, back straight, and using certain alternate fingerings for improved tone. For simplicity, let’s say that Simon played the song Jimishly, and that he played his tuba Joanishly. Again, let π be the alleged satisfier of both \( \text{PLAYED}(X, \text{SIMON, THE SONG}) \) and \( \text{PLAYED}(X, \text{SIMON, HIS TUBA}) \). Then \( \pi \) was Jimish and Joanish. We don’t need to say that \( \pi \) was \( \text{Jimish} \) for a playing of the song and \( \text{Joanish} \) for a playing of Simon’s tuba. Moreover, I don’t think we should say this.

It’s not clear what the difference is, or could be, between \( \pi \) being Jimish—in D-major, lilting, etc.—and being Jimish for a playing of the song. We can say that the logical form of (84) is (84a), without recasting (84a) as (84b).

(84) Simon played the song in the way his Uncle Jim taught him.

(84a) \( \exists x [\text{AGENT}(X, \text{SIMON}) \land \text{PLAYED}(X, \text{THE SONG}) \land \text{JIMISHLY}(E)] \)

(84b) \( \exists x [\text{AGENT}(X, \text{SIMON}) \land \text{PLAYED}(X, \text{THE SONG}) \land \text{JIMISHLY-FOR}(X, _{\text{E}})] \)

Likewise, we can say that the logical form of (85) is (85a), without recasting (85a) as (85b).

(85) Simon played the song in the way his Aunt Joan taught him.

(85a) \( \exists x [\text{AGENT}(X, \text{SIMON}) \land \text{PLAYED}(X, \text{HIS TUBA}) \land \text{JOANISHLY}(X)] \)

(85b) \( \exists x [\text{AGENT}(X, \text{SIMON}) \land \text{PLAYED}(X, \text{HIS TUBA}) \land \text{JOANISHLY-FOR}(X, _{\text{E}})] \)
The supplementary analyses of JIMISHLY(x) and JOANISHLY(x) seem unneeded and unwanted, except as side-effects of insulating Davidsonian event analyses from certain objections.

Of course, if \( \pi \) satisfies \( \text{played}(x, \text{Simon}, \text{TheSong}) \) and \( \text{played}(x, \text{Simon}, \text{HisTuba}) \) and JIMISH(x) and JOANISH(x), then \( \pi \) satisfies \( \text{played}(x, \text{Simon}, \text{TheSong}) \) & JOANISH(x) and \( \text{played}(x, \text{Simon}, \text{HisTuba}) \) & JIMISH(x). In which case, (86a) and (87a) are true.

(86) Simon played the song in the way his Aunt Joan taught him.
(86a) \( \exists x \{ \text{agent}(x, \text{Simon}) \& \text{played}(x, \text{TheSong}) \& \text{Joanishly}(x) \} \)
(87) Simon played his tuba in the way his Uncle Jim taught him.
(87a) \( \exists x \{ \text{agent}(x, \text{Simon}) \& \text{played}(x, \text{HisTuba}) \& \text{JIMISHLY}(x) \} \)
And that predicts, wrongly, that (utterances of) the English sentences (86) and (87) are true.

This makes a relativization strategy initially attractive. Many reports of “what happened” on a stage may have the logical form \( \exists x \{ \Phi(x) \& J(x) \} \). But it may not be possible to identify any “thing that happened” as the common truth maker for (the logical forms of) the reports. This makes it tempting to posit rampant relativization, until one thinks about the implications. If we could be sure that our judgments of truth/falsity in these cases coherently reflected the degree of relativization in logical forms, then judgments concerning the truth of sentential utterances might favor a relativization strategy. But in using human languages to report and think about what happened, we may report and think only in frame-dependent ways. And while we have independent evidence for framing effects that run deep in human cognition, we have no independent evidence for the requisite rampant relativization in logical forms.

5. Schein’s Resolutions

It would be nice to just stop here. Some readers may wish I had stopped sooner. But there is another perspective to consider. Schein (2002, forthcoming) offers sophisticated event analyses according to which the relevant domain—the world in which we think and talk—includes not just description-neutral events, but also “scenes” in which these events are grammatically presented or “resolved” in particular ways. His proposed logical forms include variables for both description-neutral events and grammatically individuated resolutions. As this already suggests, Schein’s goal is not to defend thesis (D).

(D) for each human language \( H \), there is a Tarski-style theory of truth that can serve as core of an adequate theory of meaning for \( H \)

He wants an empirically defensible account of the logical forms that expressions of a human language exhibit. And he follows the logic of Davidsonian event analysis where it leads, given speakers’ judgments regarding sentences of the sort discussed here; though Schein often attends to relevant details, ignored here, concerning plural and quantificational arguments. Since his proposals regarding “two-tiered” logical forms are insightful, and one can at least imagine an advocate of (D) embracing them, let me say a little about Schein’s and how I think they related to the idea of meanings as concept construction instructions.

5.1 Enriching Forms

In terms of notation, Schein replaces Davidsonian conjuncts of the form \( \Phi(x) \) with conjuncts of the form \( \exists x \{ R(x, x') \& \Phi(x') \} \); where \( R(x, x') \) is said to be true of \( \langle \epsilon, \epsilon' \rangle \) if and only if \( \epsilon' \) is the resolution of \( \epsilon \) in some scene. For example, \( \text{agent}(x, \text{AL}) \) is replaced with \( \exists x \{ R(x, x') \& \text{agent}(x', \text{AL}) \} \); where \( R(x, x') \) is true of \( \langle \epsilon, \epsilon' \rangle \) iff for some scene \( \alpha, \epsilon' \) is the resolution in \( \alpha \) of the Davidsonian (description-neutral) event \( \epsilon \), and \( \text{agent}(x', \text{AL}) \) is true of \( \langle \epsilon' \rangle \), Alvin> iff Alvin is presented as the agent in \( \epsilon' \). In effect, this adds a variable for modes of presentation of events.

With regard to (1) and (3), renumbered here as (88) and (89),
(88) Al chased Theo.
(89) Theo chased Al.
Schein would replace (thematically elaborated) logical forms like (88a) and (89a)

\[(88a) \exists x [\text{AGENT}(X, AL) \& \text{PAST}(X) \& \text{CHASE}(X) \& \text{PATIENT}(X, THEO)] \]

\[(89a) \exists x [\text{AGENT}(X, THEO) \& \text{PAST}(X) \& \text{CHASE}(X) \& \text{PATIENT}(X, AL)] \]

with variants that imply (88b) and (89b), respectively.

\[(88b) \exists x [\exists x' [R(X, x') \& \text{AGENT}(x', AL)] \& \exists x [R(X, x') \& \text{PATIENT}(x', THEO)]] \]

\[(89b) \exists x [\exists x' [R(X, x') \& \text{AGENT}(x', THEO)] \& \exists x [R(X, x') \& \text{PATIENT}(x', AL)]] \]

The one chase can be the relevant value of the matrix variable \(x\) in both (88b) and (89b). But on Schein’s view, this event can be resolved (or “redrawn”) in many ways, including ways that correspond to seeing the chase as an event that has a certain agent, or a certain patient. The hypothesized resolutions of the chase are the relevant values of the indexed variables. Each chipmunk can be seen as either an agent or a patient. But (88) presents Al as the agent and Theo as the patient, while (89) presents Theo as the agent and Al as the patient. So once \(\text{CHASE}(X)\) is replaced with \(\exists x' [R(X, x') \& \text{CHASE}(x')]\), it is clear that the verb phrases—‘chased Theodore’ and ‘chased Al’—can be consistently modified with, respectively, ‘gleefully and athletically but not skillfully’ and ‘gleelessly and unathetically but skillfully’.

More generally, an instance of (90) can be replaced with a two-tiered logical form

\[(90) \exists x [\text{CHASED}(X, AL, THEO) \& \Phi(x)] \]

that implies the corresponding instance (91), along with (88b);

\[(91) \exists x [\exists x' [R(X, x') \& \text{CHASE}(x')]] \& \exists x [R(X, x') \& \Phi(x')]] \]

where (91) might be simplified, for many purposes, as (91a).

\[(91a) \exists x \exists x' [R(X, x') \& \text{CHASE}(x') \& \Phi(x')] \]

Likewise, for (92), (89b) and (93/93a).

\[(92) \exists x [\text{CHASED}(X, THEO, AL) \& \neg \Phi(x)] \]

\[(93) \exists x [\exists x' [R(X, x') \& \text{CHASE}(x')]] \& \exists x [R(X, x') \& \neg \Phi(x')]] \]

\[(93a) \exists x \exists x' [R(X, x') \& \text{CHASE}(x') \& \neg \Phi(x')] \]

The idea is that the one chase can be simultaneously seen as a \(\Phi\)-ish chase and something Alvin did. The chase can also be simultaneously seen as a \(\neg \Phi\)-ish chase and something Theodore did. With regard to Simon’s performance, the idea is that a played-the-song resolution of that performance differs from played-his-tuba resolution; where the former resolution is telic, supporting modifiers like “in an hour”, while the latter is not.

### 5.2 Adding Contents to Meanings
I cannot here do full justice to Schein’s discussion of various puzzle cases, and the implications of taking telicity considerations seriously. But much of his discussion is devoted to arguing, quite compellingly, against other Davidsonian responses to the puzzle cases. In particular, I agree that his two-tiered approach is better than invoking rampant relativization, at least for purposes of providing logical forms that reflect the ways in which competent speakers understand sentences. Indeed, my own views have been deeply influenced by Schein’s work. But for purposes of providing logical forms that specify alleged truth conditions of human sentences, I don’t think Schein’s two-tiered approach is significantly better than positing lots of intimately related events that share their participants and a spatiotemporal address; see section four above.

My suspicion is that Schein’s proposed logical forms reflect both linguistic meanings—i.e., concept construction instructions that are generated by a human I-language and often executed via other cognitive systems—and the models of the human concepts that are assembled by executing these instructions. In my view, both tiers of his proposal are insightful, but it is still
important to distinguish linguistic blueprints from assembled concepts. That said, the difference between my view and his may be more terminological than substantive. We are both primarily concerned with the natural phenomenon of linguistic understanding; and while Schein characterizes understanding in terms of truth, the truth conditions he specifies are overly perspectival and description-sensitive in ways at odds with Davidson’s program.\textsuperscript{18}

In any case, one wants to know what the posited scene-relative resolutions are. Schein starts with an animating example. The event (or state) of Carnegie Deli facing Carnegie Hall is the event of Carnegie Hall facing Carnegie Deli. But whichever way one heads on Seventh Avenue in Manhattan, there is a left side and right side. So one can speak of a northward scene and a distinct southward scene such that everything on the left side of the former is on the right side of the latter. Though as Schein notes, other examples will strain any simple dyadic notion of scenes as perspectives on events. For example, one can weigh a car by weighing its parts. Like Davidson, Schein does not want to posit an event of weighing the car that is distinct from the co-located event of weighing the car’s parts. But here, appeal to spatiotemporal reference frames seems irrelevant. So drawing on the idea that each scene comes with a degree of resolution—think about zooming in or out with a lens—Schein suggests that every scene s comes with a “reticule” that resolves the event that s is a scene of.

The idea is that if w is the worldly event of weighing, there is a complete-car scene c and a car-parts scene p such that: Resolves(c, w); Resolves(p, w); but c and p resolve w differently. But note that even if the world includes many perspectives on each event, with perspectives individuated a Fregean way, that doesn’t yet get us resolutions as domain entities. The posited scenes have to be suitably loaded with reticules. In this sense, Schein’s appeal to scenes is not innocuous. The ontology is tendentious.

This matters, and not just for metaphysicians. On Schein’s view, (88) and (89) are false

(88) Al chased Theo.

(89) Theo chased Al.

if the world includes the two chipmunks who chased each other, and the event of them chasing each other, but not the requisite resolutions of the chase. Put another way, his theory implies that these sentences (logically) imply that there are resolutions. That’s a tendentious claim about sentences of a human I-language. One might instead view the formal sentences as proposed theoretical representations of both the language-independent events/participants that speakers think and talk about, and some ways in which speakers represent those spatiotemporally located particulars. For certain purposes, it may be legitimate to mix ontology and psychology in this way. Theorists may want to say both that speakers use sentences like (88) and (89) to assemble thoughts of a certain sort and that if all goes well, the assembled thoughts will be true or false depending on whether or not an event of a certain sort occurred. But absent independent grammatical evidence for scene variables, I don’t think that speakers tacitly quantify over scenes and resolutions, as if they are (qua competent speakers) cognizant of how complicated word-

\textsuperscript{18} In personal correspondence (and conversation), Schein says that he finds theses (D) and (L)—see section one—untenable, and that he does not identify meanings with truth conditions. He is, though, inclined to retain a modified version of (L): if S is a sentence of a human language, and S* is the logical form of S, then S has no truth condition that deviates from the truth condition if any that S* specifies. This is because Schein takes a logical form to represent both subjective and objective aspects of a speaker’s situation—her mental state and ambient conditions, at a moment of utterance—in a way the capture certain invariance(s) across expressions, thoughts, and contexts. See Ludlow (2011) for related discussion.
mind-world relations are. My inclination is to say that Schein’s resolutions are effectively the mental addresses from which concepts are called when meanings/instructions are executed.

That’s not yet an argument against Schein, as opposed to an alternative diagnosis of his conclusions. And again, sorting out the differences between views may not be easy. (Though if theories that assign truth conditions to human language sentences can be recast as theories that don’t, that’s already significant.) Like Schein and many others, I think speakers understand sentences of a human language by virtue of relating these sentences to representations of events and their participants. If one also wants to say that these sentences have truth conditions, then given the puzzle cases, the difficulties that beset other responses make Schein’s response relatively attractive and perhaps unavoidable: specify the truth conditions in terms of abstracta—viz., resolutions of events—that reflect the representations. But perhaps sentences are tools with which we often make truth evaluable claims, in contexts where all goes well, and not expressions that have truth conditions relative to contexts. Given this Strawsonian view developed in chapter four, perhaps we can make do with less intellectualized logical forms, while remembering that there are many ways in which things can fail to go well. In particular, the very linguistic capacities that let us generate sentences may render us subject to deep framing effects.
Chapter Four: Truth or Understanding

In chapter two, I discussed a model language that has a Tarskian semantics, and how such a language might be used to introduce a Frege-Church language whose expressions denote entities and truth values. But in chapter zero, I also sketched some initial reasons for doubting that there are theories of truth for the spoken/signed languages that children naturally acquire, much less truth theories that can serve as adequate theories of meaning for human I-languages. In chapter three, I focused on one cluster of reasons concerning “event” variables. In this chapter, I discuss a cluster of more general reasons for rejecting thesis (D).

(D) for each human language $H$, there is a theory of truth that is the core of a correct theory of meaning for $H$

1. Overview

Some reasons for rejecting (D) are independent of any particular conception of what human languages are. Though given thesis (C), discussed in chapter zero,

(C) Each human language is an I-language in Chomsky’s sense.

(D) implies the very bold conjecture (B).

(B) Each human language is a biologically implementable generative procedure $H$ such that some theory of truth for $H$

is the core of a correct theory of meaning for $H$.

In my view, (B) is very implausible, and this tells against (D). While truth-theoretic models of human I-languages can be valuable, the interpretations that such models connect with bracketed strings of lexical items—e.g., [see [a [man [with [a telescope]]]]]—seem to have very different properties than the meanings that child-acquirable procedures connect with pronunciations in certain human ways. This first section is an outline of the argument developed below.

1.1 Liars

In section two, I discuss “Liar Sentences” like (1);

(1) Lari is not true.

where ‘Lari’ is, by stipulation, a name for the second numbered sentence in this chapter.¹

(2) The second numbered sentence in chapter four of Conjoining Meanings: Semantics Without Truth Values is not true.

I think that (1), (2), and their kin vividly illustrate a general property of human language sentences: they are neither true nor false—and not because they have other truth values, or have truth values relative to contexts. In my view, sentences of a human I-language are like most things in not having truth conditions. In this respect, (1-3) are like galaxies, horses, and numbers.

(3) Snow is white.

Such things are neither true nor false because they are not truth-evaluable. Likewise, in my view, neither (4) nor (5) has a truth condition; and hence, neither of these I-language sentences are true.

(4) The sentence ‘Lari is not true.’ is true if and only if Lari is not true.

(5) The sentence ‘Snow is white.’ is true if and only if snow is white.

I readily grant that there are truth-evaluable things, including many acts of making judgments or assertions by using sentences like (3). And it may be that Liar Sentences cannot be so used. But this doesn’t show that sentences like (1) or (2) have defective meanings. The corresponding thoughts may well be non-ideal in various respects. So for various purposes, we

¹ Compare the introduction of ‘Julius’ as a name for whoever invented the zip; see Evans (1982). For many purposes, we could use ‘(2)’ instead of ‘Lari’. But ‘(2)’ is not a human language name for sentence (2); and ‘(2) is not true’ may not be an expression of any human I-language, as opposed to semi-technical analog.
may want to speak of propositions when talking in idealized ways about the truth-evaluable contents of judgments/assertions. Though a decision to talk this way is not an argument for the hypothesis that sentences of a human I-language determine (function from contexts to) propositions. On the contrary, sentences like (1) and (2) suggest that this hypothesis is false, even if we agree to waive concerns about which proposition is allegedly expressed by (6)

(6) Al chased Theo.
in contexts where the possibility of mutual chase has not been excluded. And whatever we say about judgments/assertions/propositions, Liar Sentences tell against thesis (D1).

(D1) for each human language H, there is a correct theory of truth for H.
There are potential replies. But as we’ll see, these replies make it harder to defend (D).

(D) for each human language H, there is a theory of truth
that is the core of a correct theory of meaning for H

Employing certain technical notions, familiar to logicians, can help in providing truth-theoretic models that accommodate analogs of (1) and (2). But especially given thesis (C),

(C) Each human language is an I-language in Chomsky’s sense.

these technical notions seem out of place in theories of meaning for child-acquirable languages.

It can be tempting to just set Liar Sentences aside, along with talk of skies and sunrises, as special cases. But if human languages are biologically implemented procedures that generate expressions—then we shouldn’t be surprised if the generable expressions are not related to the world in ways required for truth-evaluable. Sentences like (3) may be tools that speakers use, in assembling thoughts from available concepts, rather than end-products that themselves have truth values (even relative to contexts). If (3) is a biologically generable combination of lexical items, which humans can use to assemble a thought of a certain sort, then (3) may not be any more truth-evaluable than (1). It may just be easier to see that (1) is not truth-evaluable.

1.2 Contexts
It is worth remembering that even sincere assertions can fail to be true or false.

A competent speaker might sincerely utter (2)

(2) The second numbered sentence in chapter four of

Conjoining Meanings: Semantics Without Truth Values is not true.

while justifiably but mistakenly thinking that (7)

(7) 2 + 2 = 5

was the second numbered sentence in the relevant chapter. Such a speaker might also utter (8),

(8) Vulcan has an iron core.

presupposing that Vulcan exists in order to say that it is dense enough to affect Mercury’s orbit. No such utterance is true. But falsity requires more than grammaticality, sincerity, and nontruth; see Strawson (1950), Evans (1982). Likewise, an utterance of (9) may be neither true nor false

(9) That is a dog.

if the utterer fails to demonstrate anything. More generally, making a false claim requires a significant kind of success, in that various presuppositions have to be met.2

In section three, I argue that these familiar points bolster “Liar Arguments” against (D),

(D) for each human language H, there is a theory of truth
that is the core of a correct theory of meaning for H

---

2 Cp. Armour-Garb (2001, 2004) on what he calls the “no proposition expressed” strategy. Though it is hard to stop (2) from expressing a proposition if (2) would have expressed a proposition had (7) been the second numbered sentence, unless the relevant notion of expressing is technical and presupposes an independent sense in which (2) is meaningful. See Armour-Garb and Woodbridge (2013, forthcoming).
in part by highlighting the question of whether a theory of truth/meaning for a human language should have theorems concerning utterances, sentences, or both. This foundational question is interwoven with questions about whether human I-languages generate expressions with elements that track the many ways in which the truth or falsity of utterances can depend on context; cp. Stanley (200x). Section three connects these sources of difficulty for (D) with others that are more specific to complementizer phrases, as in (10).

(10) Everyone who thinks that there are unicorns on Hesperus wants to join Paderewski in hunting them.

As Kripke’s (197x) puzzle highlights—and as Davidson’s (196x) initial discussion of “samesaying” already suggested—reports like (10) and the simpler (11)

(11) Someone denied that Hesperus is Phosphorus.

do not merely present the challenge of providing a model according to which replacing one name for Venus with another need not preserve truth. Even given a theory according to which ‘that’-clauses can be used to denote (or describe, or quantify over) suitably fine-grained truth-evaluable entities, there remains the question of whether words like ‘think’ and ‘deny’ have Tarski-style satisfaction conditions. Prima facie, the truth or falsity of an utterance of (11) can depend on the context in many complicated ways. So even if ‘chase’ has a satisfaction condition that can be specified, perhaps allowing for some dimensions of context sensitivity that can be plausibly accommodated in a specification of what ‘chase’ means, one might doubt that ‘think’ and ‘deny’ (‘say’, ‘doubt’, etc.) have any such satisfaction conditions.

It can be tempting set these lexical items aside, along with ‘sky’ and ‘Venice’, as special cases. But at some point, advocates of (D) have to get beyond sketching idealized models of various constructions, and provide a model that is accurate for at least a few simple cases. We don’t expect models of the hydrogen atom to extend easily to all of chemistry, much less biology. But a theory of the atomic table cannot abstract away from the rows, the columns, and each of the elements. Likewise, a truth-theoretic account of meaning for a human I-language cannot idealize away from each lexical item while also bracketing general questions about truth, denotation, predicates, relations, and contexts.

A main theme of this chapter and its predecessor is that nearly fifty years after Davidson’s (1967a, 1967b, 1968) seminal work, a cluster of objections to (D) remain as serious as they were then. Some weaker objections in the vicinity have been addressed, and various proposals have been offered, showing that the stronger objections are not decisive. But we can and should ask whether the ancillary assumptions required to defend (D) are plausible. For truth-theoretic semantics is neither in its infancy nor the only game in town when it comes to describing the meanings of human language expressions. As discussed in chapter three, we can reject Davidson’s conception of meaning, but welcome his replacement of CHASED(X, X’) with CHASED(X, X’, X’’) in logical forms. Given a hypothesized logical form L for a human language sentence S, we can view L as a model thought that might be constructed from ideal atomic concepts by executing the assembly instruction that S provides, instead of saying that L specifies the truth condition that S itself exhibits. On this view, the meaning of S does not determine a truth condition for S. So we need not abandon all the good work on action reports, Liar Sentences, context sensitivity, and the meanings of complementizer phrases. But we need to locate the real value of this work, as opposed to automatically construing it in terms of (D).

2. Liars, Truth, and Meaning

In §2.1, I discuss the difficulty that Liar Sentences present for (D1).

(D1) for each human language H, there is a correct theory of truth for H.
I then turn to an independent worry—via Foster (1976)—about (D).

(D) for each human language H, there is a theory of truth
that is the core of a correct theory of meaning for H

As discussed in §2.2, a certain response to “Foster’s Problem” has become standard. But this response makes it even harder to see how advocates of (D) can accommodate examples like (1)

(1) Lari is not true.

2.1 Ways of Not Being True
Recall that Lari is none other than sentence (2).

(2) The second numbered sentence in chapter four of
Conjoining Meanings: Semantics Without Truth Values is not true.

The basic point regarding such sentences is familiar. If (2) is true, then since (2) is the second numbered sentence in the relevant chapter—viz., this one—then (2) is not true. But if (2) is false, then it is false that (2)—a.k.a. Lari—is not true; yet if (2) is false, then given how falsity related to truth, (2) is not true. So if (2) is true, then (2) is true, and (2) is not true; and if (2) is false, then it is false that (2) is not true, and (2) is not true. So (2) is neither true nor false. Likewise, (1) is neither true nor false. Hence, Lari is not true, and (1) is not true.

That conclusion is not paradoxical. Many things are neither true nor false. Consider the galaxy Andromeda, my horse Sadie, the number 5, etc. If (1) and (2) are among the things that are neither true nor false, that is not deeply puzzling. If (1) and (2) are among the things that are true nor false, then we face puzzles; but this is one reason, among many, for thinking that (1) and (2) are not among the things that are true nor false. There is nothing paradoxical in an especially good argument that Andromeda is not true. Likewise, I claim, there is nothing paradoxical in an especially good argument that Lari is not true. Given the further assumption that (1) is true iff Lari is not true, contradiction follows; but this is one reason, among many, for thinking that the further assumption is false. There are nearby puzzles concerning truth, judgment, and assertion. But the premises of a paradox are supposed to be compelling. And we can consistently reject the hypothesis that sentences of a human language are true or false. Though before drawing morals from (1), we need to be clear that there is more than one way of being neither true nor false.

In noting that Andromeda is neither true nor false, one is not asking which other truth value(s) it exhibits. Many things are just not truth-evaluable. In certain contexts, one might be able to use (12) as an indirect way of saying that Andromeda is not true.

(12) Andromeda is true if and only if there are finitely many prime numbers.

But absent the stipulations of a special invented language, no galaxy has a truth condition. Galaxies, unlike sentences of PL, do not have satisfaction conditions. Nor do galaxies express thoughts and thereby inherit truth conditions. Similar remarks apply to horses, numbers, etc.

In my view, sentences of a human language fail to be true or false in this boring sense. They are not among the truth-evaluable things.

We can invent a language, ηPL, such that: its sentences are truth-evaluable, and a sentence is false if and only if its negation is true; it allows for symbols of type <> that do not stand for any d-entity; and sentences of ηPL that contain such symbols denote a third truth value η, distinct from T or ⊥. In various ways, one can also stipulate that the invented sentence (13)

(13) False(Triskaideka)

Likewise, each sentence/galaxy/horse is neither greater than 5 nor less than or equal to it, and yet—unlike 5i—not a special kind of number. Say, if you like, that some true claims are contradictory; see Priest (1979, 2006). We still need very good reasons for thinking that each one of some contradictory claims is true. And in my view, the reasons for thinking that human language sentences are truth-evaluable do not motivate acceptance of contradictions.
denotes η if “Triskaideka” denotes (13); cp. Kleene (1950), Kripke (1975). One can then hypothesize that human languages are like ηPL in this respect. On this view, (14)

(14) The fourteenth numbered sentence in chapter four of
Conjoining Meanings: Semantics Without Truth Values is false.

is relevantly like (13). So if ‘Aril’ is a name for the fourteenth numbered example in this chapter, then perhaps (15) has the truth condition specified by (15-T), both sides of which are false.

(15) Aril is false.

(15-T) True(‘Aril is false.’) ≡ False(Aril)

In which case, both (15) and Andromeda fail to be true or false, but (15) has the truth value η.

So one can coherently hypothesize that (15) is true if and only if Aril is false. One can also say that a homophonic/homographic analog of (15) would have been true, had the fourteenth numbered example in this chapter been sentence (7), repeated below.

(7) 2 + 2 = 5

And one speculate that (8) has a truth condition, or least a truth value, without being true or false.

(8) Vulcan has an iron core.

But whatever the attractions of this idea, not mine, (1) does not merely fail to be true or false.

(1) Lari is not true.

For sake of argument, let’s grant that sentences can have the truth value η. If (1) is like (13) in being truth-evaluable, then presumably, (1) is true if Lari is not true; and Lari is not true, even if there is a difference between being false and being truth-evaluable but not true. So if (1) is truth-evaluable, then presumably, (1) is true. But (1) isn’t true. Perhaps (1) is truth-evaluable in some technical sense without having the truth condition specified by (1-T).

(1-T) True(‘Lari is not true.’) ≡ ~True(Lari)

But it needs showing that any such technical sense has anything to do with linguistic meaning. And note that even if sentences can have the truth value η, (1-T) is not true; its right side is true, but its left side isn’t. So (1) seems to be among the things that are not truth-evaluable, and likewise for boundlessly many other Liar Sentences.

I suspect that if (1) is not truth-evaluable, then neither is (15), even if we cannot prove that (15) doesn’t have the value η. Logicians have reasons for distinguishing “mere” Liar Sentences like (14) from their “strengthened” counterparts that replace ‘false’ with ‘not true’. But it’s hard to see how a human I-language could respect this relatively subtle distinction, somehow generating expressions in a way that leads to (15) having a truth condition—and perhaps the truth value η—while (1) ends up not being truth-evaluable. Though likewise, it’s hard to see how (16-18) could be truth-evaluable if (1) and (15) are not; see §2.4 below.

(16) Lari is a sentence.

(17) Bessie is not blue.

(18) Bessie is a brown cow.

So perhaps no human language sentences are truth-evaluable, and Liar Sentences are just especially vivid examples, as opposed to sentences that call for special semantic treatment.

Of course, it’s possible that only some declarative sentences have truth conditions. One can speculate that the declaratives of a human I-language have truth conditions unless they are defective in a way that (1) is. But absent a good account of how this possibility is realized, it remains a mere speculation. If sentences like (1) were the only examples of declaratives that seem not to have truth conditions, then it might be reasonable to set them aside as special cases. But as we’ve seen, doubts arise even with regard to apparently humdrum cases like (6);

(6) Al chased Theo.
and (3), while famous, is hardly a good example for Davidsonians.

(3) Snow is white.

Still, it’s widely held that modulo miscreants, human language sentences are true or false relative to contexts. This leaves multiple degrees of freedom with regard to accommodating Liar Sentences: either the sentences themselves or the contexts in which they are used are somehow exceptional in a way that can be ignored when specifying the meanings of expressions like (3) and (6), which do not include “semantic” vocabulary like “true’. Arguing against this vague response to Liar Sentences will take a while. Though in the end, I think the basic point is about as simple as it seems to be at first glance: (1) has a perfectly fine meaning, but no truth condition;

(1) Lari is not true.

and there is no relevant difference between (1) and (17), at least not for purposes of theorizing about human languages and the meanings they connect with pronunciations.

This is not to deny that context matters with regard to examples like (19).

(19) That is not true.

Logicians can use (19) to illustrate a paradox concerning the assertions that speakers can make in certain contexts. Some theorists will want to say that such contexts are somehow defective, and that (19) has a truth value relative to each nondefective context. I disagree; but this is a conclusion, not a premise. I readily grant that (19) can be used to make claims about many things, including (1-20), and that many of these claims are true or false.

(20) 2 + 2 = 4

And one can talk about relations that a sentence S bears to the pairs (v, C) such that v is the truth value of the assertion that a speaker would make in the non-defective context C by using S to make an assertion. I don’t think (19) has a truth value relative any to any context. But my claim is not that (19) fails to have a truth condition because of Liar Paradoxes, and that (1) is similar. My claim—which does not rely on special assumptions about truth, assertions, or contexts—is that (1) does not have a truth condition. If this is correct, then advocates of (D) face the challenge of providing a plausible theory according to which (17) has a truth condition but (1) does not.

(17) Bessie is not blue.

Again, one can invent languages whose sentences include analogs of (1); where truth conditions for these sentences are stipulated in terms of technical notions like Kripke’s (1975) fixed points, or Gupta and Belnap’s (1993) revisions. Such a language is related to the relevant domain in a logically sophisticated way, as specified by the inventor. But for just this reason, the inventions may be unlike human languages. So even if certain natural languages of thought are biological variants of (C)PL—with open sentences like ABOVE(X, X’) and COW(X), as discussed in chapter two—it hardly follows that some such mentalese is related to the world in the way characterized by certain technical notions that logicians designed with the aim of consistently assigning truth conditions to invented analogs of (1).4

---

4 As discussed in chapter one, it’s hard enough to explain how any animals could have a mental variant of (C)PL, even one that doesn’t support logically sophisticated truth theories. A concept may not be true of anything unless the concept is sufficiently integrated with others in ways that make it possible for the concept-user to recognize logical relations among concepts that are sufficiently numerous and diverse. Achieving such integration is no small task. And it’s hard to see how words could be true of things if not via concepts. It’s puzzling enough that COW(_) applies to all cows, but nothing else. Any corresponding relation that ‘cow’ bears to the cows would seem to be more complicated. Children apparently use COW(_) to acquire ‘cow’; see Bloom (200x). And as Fodor (1986, 2003) stresses, echoing Chomsky (1959), the presence of cows often elicits uses of COW(_) but not ‘cow’; though cows
Tarski (1933, 1944) had his own reasons for denying that there are theories of truth for human languages. But even if these reasons can be evaded by quarantining sentences like (1), we shouldn’t evade the question of whether sentences like (17) are truth-evaluable. As many philosophers have stressed, it’s remarkable that humans can make any claims that are true or false. Once a thinker achieves this level of clarity and contact with her environment, truth is at most a mere negation sign away. If only for this reason, it is very optimistic to think that sentences of a human language—expressions that children can generate and comprehend—are themselves truth-evaluable. In the right settings, humans can use these sentences to make and express truth-evaluable judgments, given suitable concepts. But truth/falsity may be downstream of linguistic meaning, in that certain acts of using meaningful expressions are candidates for being true or false, subject to further constraints that are not specifically linguistic.

2.2 Same Truth Value, Different Meanings
In thinking about the implications of Liar Sentences for (D),

(D) For each human language H, there is a theory of truth that is the core of correct theory of meaning for H.

we need to think about the ways in which (D) goes beyond (D1).

(D1) For each human language H, there is a correct theory of truth for H.

In particular, if a theory of meaning for a human language is a theory of understanding for that language, one might wonder if a Tarski-style theory of truth that accommodates (11)

(1) Lari is not true.

can also be a theory of understanding for English. But let’s go slowly and try to be concessive.

If we abstract away from any context sensitive elements of (21) and (22),

(21) Ernie snores.
(22) Bert yells.

then a Tarski-style theory of truth for English—or better, for any version of English that generates analogs of (21) and (22)—will have theorems like (21-T) and (22-T).

(21-T) True(‘Ernie snores.’) ≡ Snores(Ernie)
(22-T) True(‘Bert yells.’) ≡ Yells(Bert)

In fact, (21) is complicated in at least three ways. First, ‘snores’ has a generic character that ‘snored’ does not. (Tense introduces its own complexities, as does ‘is a snorer’.) Second, there are many Ernies; see, e.g., Burge (1973). Third, ‘Ernie’ is being used here to talk about a muppet. And in many contexts, it is clear that muppets are silent inanimate things that cannot snore. Yet it doesn’t follow that (21) is false; cp. Travis (1985). Similar points apply to (22) and other instances of the ‘NAME VERB’ form. But let’s pretend that (21-T) and (22-T) are actually true, with invented sentences like ‘Snores(Ernie)’ having suitable interpretations by stipulation.

Once we consider more complex sentences, or homophony, it becomes clear that for purposes of deriving T-sentences, mere quotation is inadequate as a device for depicting object language sentences. We need structural descriptions like [see [the [man [with [a telescope]]]]]. But let’s view quotation as standing in for suitable structural descriptions of the intended expressions; though it may turn out that by appealing to such descriptions, defenders of (D) incur a commitment to a conception of human language syntax that presupposes (C).

(C) each human language H is an I-language in Chomsky’s sense

Let’s also recognize that (D) does not imply that every theory of truth is the core of a correct theory of meaning. Following Davidson (1967b, 1984), one might say that a meaning rarely if ever elicit ‘cow’ absent uses of Cow(_). So if the word is true of what the concept applies to, then presumably, this semantic relation is conceptually mediated.
theory requires a suitably formulated truth theory that meets certain empirical constraints. Indeed, there are at least two clusters of reasons for saying this. I’ll soon return to the second. But first, and perhaps most obviously, human languages allow for nondeclarative sentences. A theory of truth/meaning for English needs to be formulated in a way that makes it possible to supplement the theory in ways that accommodate (23-28).

(23) Ernie snores?  (24) Does Bert yell?
(25) When does Ernie snore?  (26) Does Bert yell when Ernie snores?
(27) Please color Kermit green.  (28) Don’t color Kermit red!

One obvious thought, developed in various ways by many authors, is that each sentence inflects a sentential “radical” with a grammatical mood; where the radical has a truth condition, perhaps relative to assignments of values to variables, and the mood defeasibly indicates the kind of speech act being performed. Abstracting from context sensitivity, the idea is that (21) and (23) share a part that is true if and only if Ernie snores, while the difference in pronunciation reflects a difference in mood that is meaningful, at least with regard to how acts of uttering these sentences are construed. Some theories (e.g., Hamblin 1958, 1973) go farther, treating interrogatives and declaratives as expressions that denote things of different sorts—e.g., questions and propositions. It may be that an adequate truth-theoretic conception of linguistic meaning will need to take some such form, in order to describe the systematic relations exhibited by matrix questions like (25), relative clauses as in (26), and embedded questions as in (30).

(30) Does Bert know when/where/why/if/whether Ernie snores?

But let’s not fuss about this. Given a theory of truth that can serve as a decent theory of meaning for English declaratives, it would be churlish to complain if the theory didn’t explain everything we would like to explain about the meanings of nondeclarative sentences.

However, a correct theory of meaning for a “declarative fragment” of a human language H has to be formulated in a way that permits a correct extension of the theory to the unboundedly many interrogative and imperative sentences of H. So while (D) does not imply that every truth theory for H can be supplemented in ways that accommodate the nondeclarative sentences of H, (D) does imply that some truth theory for H is so supplementable. It’s not enough if some theory of truth might be supplemented, in the sense that there is no barrier in principle to using a truth theory as the core of a theory of meaning. Empirically discovered features of human languages may preclude extending a truth theory for declaratives to nondeclaratives. I won’t press such objections here, since the grammatical details are complicated, and any proposal is sure to be controversial; see Lohndal & Pietroski (2011). But objections to (D) need not establish that either (D1) is false, or a truth theory cannot be a meaning theory.

(D1) for each human language H, there is a correct theory of truth for H.

This matters, because objections to (D) take various forms. Some objections, regarding specific constructions, target (D1) without challenging the idea that such a theory (if there is one) could serve as a theory of meaning. Some objections to this latter idea, concerning how meaning is related to truth, are compatible with there being theories of truth for human languages. But objections to (D) can also take a more dilemmatic form: given certain features of H, no plausible truth theory for H can be formulated in a way that makes it plausible meaning theory. In

---

5 See Dummett (1976) on the need for a theory of “content” that meshes with a theory of “force,” whether or not ‘True(_)_’ is the central predicate in a theory of meaning for a Human Language; cp. Wright (197x). See also Segal (1991), Ludwig & Lepore (200x), Lohndal & Pietroski (2011), and further references there.

6 Even a kid from Kenosha can become the president. But it doesn’t follow that some kid from Kenosha will be elected if suitably formulated and supplemented with campaign funding.
particular, a truth theory that accommodates Liar Sentences might be too sophisticated to be a meaning theory for a language that a child can naturally acquire. Though to develop this thought, we need to consider a general concern about using truth theories as meaning theories. And this brings us to the second reason for thinking that only suitably formulated truth theories that meet certain empirical constraints can serve as theories of meaning for human languages.

Sentences (21) and (22) are not synonymous. But they may be alike with regard to truth.

(21) Ernie snores.
(22) Bert yells.

Suppose that Ernie snores and Bert yells, and hence that the invented sentence (31) is true,

(31) Snores(Ernie) & Yells(Bert)

along with the T-sentences (21-T) and (22-T). Then (32-34) are also true.

(21-T) True(‘Ernie snores.’) ≡ Snores(Ernie)
(22-T) True(‘Bert yells.’) ≡ Yells(Bert)
(32) True(‘Ernie snores.’) ≡ Yells(Bert)
(33) True(‘Bert yells.’) ≡ Snores(Ernie)
(34) True(‘Ernie snores.’) ≡ True(‘Bert yells.’)

Similarly, since (35a) and (35b) both fail to be true, (35c) is true.

(35a) There are finitely many primes.
(35b) Five precedes two.
(35c) True(‘There are finitely many primes.’) ≡ True(‘Five precedes two.’)

Indeed, endlessly many instances of ‘True(S) ≡ p’ are true but “uninterpretive” in the following sense: the metalanguage sentence used on the right does not specify the meaning of the object language sentence mentioned on the left. So one wants to know how a theory of truth could ever be a theory of meaning for a language whose sentences have meanings that are individuated more finely than truth values. There may be possible worlds at which (32-34) are false, holding the meanings of (21) and (22) fixed. But given endlessly many examples like (35), sentence meanings seem to be more finely individuated than sets of possible worlds.7

In response, an obvious initial thought is that one can construct a theory whose theorems include (21-T) and (22-T), but not (32-34). Suppose there is a truth theory Θ for English such that no uninterpretive T-sentences can be derived from Θ. Then Θ might serve as the core of a meaning theory for English, since Θ would—via its theorems—pair each sentence S of English with (and only with) a sentence of the metalanguage that specifies what S means. But this line of thought raises at least three clusters of questions that advocates of (D) must address.

First, are T-sentences like (21-T) meaning-specifying in the relevant sense, or do “interpretive” T-sentences need to be more disquotational and hence more like (36)?

(36) ‘Ernie snores.’ is true if and only if Ernie snores.

If the latter, then it seems that no interpretive T-sentences will be derivable, given the need for at least some technicality in the metalanguage. This point becomes especially vivid once we consider context sensitivity and quantification, as illustrated with (37); see section three.

(37) I chased something.

But let’s grant that T-sentences like (21-T) are interpretive. More generally, let’s not fuss if the right sides of derivable T-sentences exhibit logical structure, so long as they don’t employ notions that are “foreign” to the sentences mentioned (or structurally described) on the left, as in

---

7 Given any independently plausible notion of possible worlds; see Kripke (1982), Lewis (1986).
(32-34); though the interpretive/uninterpretive (or domestic/foreign) distinction may have to be drawn in terms of the cognitive resources that humans employ when acquiring using languages.

Second, there is the question of whether any theory has enough deductive power to generate interpretive T-theorems, yet not enough to generate uninterpretive T-theorems. The issue here is not merely the technical one of writing appropriate axioms and inference rules. Tarski envisioned a theory of truth as an addition to a nonsemantic theory (e.g., of arithmetic) that is formulated in an object language governed by a logic that licences deduction of theorems concerning the domain (not just the object language). One can instead view a hypothesized truth theory for English as a theory in its own right, *but not* part of a larger theory-of-the-world that is formulated in the same metalanguage and subject to the same logic. I think it’s odd to view truth theories as segregated in this way from theories of the things that words are allegedly true of. But let’s grant that a theory of truth for English need not be interwoven with assumptions that would make it possible to derive (38) from (21-T).

\[
(38) \text{True(‘Ernie snores.’)} = \text{Snores(Ernie)} \& \text{Precedes(Two, Five)} \\
(21-T) \text{True(‘Ernie snores.’)} = \text{Snores(Ernie)}
\]

This still leaves the question of whether any theory \( \Theta \) that generates theorems like (21-T) can fail to generate uninterpretive theorems like (39a) or (39b); where \( \Gamma \) is itself a theorem of \( \Theta \).

\[
(39a) \text{True(‘Ernie snores.’)} \equiv \text{Snores(Ernie)} \& \Theta \\
(39b) \text{True(‘Ernie snores.’)} \equiv \Gamma \supset \text{Snores(Ernie)}
\]

The scope of this point depends on the relevant background logic. But consider a Fregean (second-order) logic that licenses instances of (40); where \( A \) is a theorem of arithmetic,

\[
(40) \text{HP} \supset A
\]

and ‘HP’ stands for (the conjunction of suitable Fregean definitions and) the Humean Principle that some things correspond one-to-one with some things if and only if the former have the same cardinality as the latter; see Heck (2011) for extended discussion. If instances of (40) can be instances of \( \Gamma \) in (39a) or (39b), then generating uninterpretive T-sentences is easy. One can, of course, try to formulate theories of truth against the background of a weaker logic. But a truth theory that has (21-T) and (22-T) as theorems will also have (4) as a theorem

\[
(17-T) \text{True(‘Bert yells.’)} = \text{Yells(Bert)} \\
(41) \text{True(‘Ernie snores.’)} \equiv \text{Snores(Ernie)} \& [\text{True(‘Bert yells.’)}] \equiv \text{Yells(Bert)}
\]

if the theory permits replacement of a sentence ‘p’ with the conjunction of ‘p’ and a theorem. So the background logic needs to be very weak, perhaps to the point of being a logic in name only. I return to this point in section three. But the requisite isolation of a truth theory is quite severe. Elementary principles of logic can yield a theory that generates uninterpretive T-sentences.

We are, recall, reviewing some of the questions raised by the following idea: if there is a (suitably supplementable) truth theory \( \Theta \) for English such that no uninterpretive T-sentences can be derived from \( \Theta \), then \( \Theta \) is the core of a theory of meaning for English. The first two clusters

---

8 It’s worse than odd if one also says that ‘there is no boundary between knowing a language and knowing our way around in the world’ (Davidson 1986, p. 446). It seems that (D) requires a significant boundary between general world knowledge and speaker knowledge concerning the implications of truth theories, pace Quine (195x). But for now, let’s not worry about whether advocates of (D) are entitled to isolate truth theories for human languages from theories of numbers, spacetime, matter/energy, life, animal psychology, human language syntax, etc.

9 If derivational principles apply only to premises of one sort, corresponding to one aspect of cognition—without supporting extraction of consequences from premises that reflect diverse forms of thought—why view the principles as a logic, rather than mere pattern matching within a single (perhaps computational) cognitive module?
of questions concerned the plausibility of this conditional’s antecedent. A third cluster grants the antecedent, for the sake of argument, and challenges the conditional itself.

2.3 Fostering Liars

Foster (1974) observed that given a truth theory which generates T-theorems like (21-T),

(21-T) True(‘Ernie snores.’) ≡ Snores(Ernie)

one can easily construct extensionally equivalent theories whose theorems include corresponding instances of (42), where ‘T’ can be replaced with any true sentence of the metalanguage.

(42) True(‘Ernie snores.’) ≡ Snores(Ernie) & T

For illustration, suppose that (21-T) follows from a theory that has (43) and (44) as axioms,

(43) ∀x[TrueOf(‘Ernie’, x) ≡ (x = Ernie)]
(44) ∀x[TrueOf(‘snores’, x) ≡ Snores(x)]

along with some schema like (45), whose instances include (45a).

(45) True(‘NAME VERB.’) ≡ ∃x[TrueOf(‘NAME’, x) & TrueOf(‘VERB’, x)]
(45a) True(‘Ernie snores.’) ≡ ∃x[TrueOf(‘Ernie’, x) & TrueOf(‘snores’, x)]

Replacing (44) with (46), or (45) with (47), yields a theory from which (48) follows.10

(46) ∀x[TrueOf(‘snores’, x) ≡ Snores(x) & Yells(Bert)]
(47) True(‘NAME VERB.’) ≡ ∃x[TrueOf(‘NAME’, x) & TrueOf(‘VERB’, x)] & Yells(Bert)
(48) True(‘Ernie snores.’) ≡ Snores(Ernie) & Yells(Bert)

So even if some truth theory for English has (21-T) as a theorem, this doesn’t yet explain why (21) means what it does. An alternative theory has (48) as a theorem.

(21) Ernie snores.

At least one of these truth theories is not the core of a correct theory of meaning for English, since (21) and (49) differ in meaning.

(49) Ernie snores and Bert yells.

To be sure, (46-47) are not interpretive axioms, while (44-45) may be. But it’s no defense of (D) to say that a truth theory can be a meaning theory so long its axioms are interpretive.

(D) For each human language H, there is a theory of truth that is the core of correct theory of meaning for H.

If a theory of truth is to explain why human language expressions mean what they do, then we need some account of what makes (44-45) better than (46-47), as theoretical descriptions of the semantic properties allegedly exhibited by the expressions in question.

Davidson (1984) evidently agreed, since he suggested that uninterpretive axioms fail to meet certain empirical constraints, which he tried to specify by talking about what an idealized interpreter would ascribe to speakers on the basis of certain evidence; see section four. Many authors have rightly objected to the verificationistic (and Euthyphronic) conception of meaning and language acquisition that Davidson encoded by appealing to what a “Radical Interpreter” would say.11 So let’s pass over this aspect of Davidson’s proposal. In my view, unabashed cognitivism is a far better option. One can hypothesize that speakers of H have a language of

---

10 For simplicity, I assume we can derive ‘Snores(Ernie)’ from ‘∃x[(x = Ernie) & Snores(x)]’ and ‘Snores(Ernie) & Yells(Burt)’ from ‘∃x[(x = Ernie & Snores(x)) & Yells(Burt)]’. But this is not essential to the general point.

11 See Fodor and Lepore (1992), Heck (2007), Pietroski (2005b), Ludwig and Lepore (200x). And of course, it doesn’t help to replace the Radical Interpreter with a semanticist who inscribes a theory Θ—especially if the metalanguage is as powerful as Church’s lambda calculus—absent a justification for adopting Θ as opposed to extensionally equivalent alternatives.
thought in which they encode truth-theoretic axioms, from which consequences can be extracted in a constrained way. The idea is that certain axioms like (44-45) are better than (46-47) because competent speakers *encode and deploy* the former but not the latter.

As noted in chapter two, Larson and Segal’s (1995) system for deriving T-theorems is an admirably explicit paradigm. It permits instantiation of axiomatic schemata that specify hypothesized contributions of phrasal syntax to phrasal meaning. But the rules for deriving further consequences, given axioms that specify the hypothesized contributions of lexical items, only permit replacements of established equivalents: if \( \vdash P = Q \) and \( \vdash Q = R \), then \( \vdash P = R \); if \( \vdash \alpha = \beta \), then \( \vdash \Phi \alpha = \Phi \beta \). So the inference from (21-T) to (42) is not licenced, not even if ‘T’ is itself a theorem.

\[
\begin{align*}
(21-T) & \quad \text{True(‘Ernie snores.’)} \equiv \text{Snores(Ernie)} \\
(42) & \quad \text{True(‘Ernie snores.’)} \equiv \text{Snores(Ernie)} \wedge T
\end{align*}
\]

And while theorists inscribe theorems like (21-T), the idea is that speakers can derive mental analogs like (16-M) from more basic representations; where small capitals still indicate concepts.

\[
\begin{align*}
(21-M) & \quad \text{True(‘Ernie snores.’)} \equiv \exists x [(x = \text{ERINE}) \wedge \text{SNORES}(x)] \\
(43-M) & \quad \forall x [\text{TRUE-OF(‘Ernie’, x)} \equiv (x = \text{ERNIE})] \\
(44-M) & \quad \forall x [\text{TRUE-OF(‘snores’, x)} \equiv \text{SNORES}(x)] \\
(45-M) & \quad \text{True(‘NAME VERB.’)} \equiv \exists x [\text{TRUE-OF(‘NAME’, x)} \wedge \text{TRUE-OF(‘VERB’, x)}]
\end{align*}
\]

Heim and Kratzer’s (1998) system can be spelled out in a similar way, permitting derivations of mental sentences like TRUE(‘Ernie snores.’) \( \equiv \text{SNORES(Ernie)} \) from Church-style axioms like the following: \( \ll \text{snores} \| = \lambda x. T \equiv \text{SNORES(x)} \); TRUE(‘...’) \( \equiv \ll \ldots \| = T \).

Of course, appeal to a language of thought abandons certain ambitions for theories of meaning; cp. Dummett (1975). And by focusing on how speakers actually represent the alleged truth conditions of English sentences, one goes beyond the task of specifying a theory such that knowledge of it would suffice for knowing that these sentences have those truth conditions; cp. Davidson (1967b, 1976), Foster (1976). But these are not yet objections to a cognitivist gloss of (D) and how human semantic competence can be applied to specific expressions.\(^{12}\) Some ambitions are unattainable, and some tasks are unduly modest. Advocates of (D) are free to offer other conceptions of the constraints that are allegedly satisfied by at least one truth theory for each human language, yet not by extensionally equivalent uninterpretable theories. But let’s grant, at least for now, that some constrained system of derivation provides an answer to Foster’s Challenge.

Still, an obvious worry remains. If derivability is logically blind, apart from replacement of established equivalents, then (1-T) and its mental analog will be as derivable as (13-T).

\[
\begin{align*}
(1-T) & \quad \text{True(‘Lari is not true.’)} \equiv \neg \text{True(Lari)} \\
(17-T) & \quad \text{True(‘Bessie is not blue.’)} \equiv \neg \text{Blue(Bessie)}
\end{align*}
\]

\(^{12}\) And given such a gloss, theorists can say that a hypothesized truth theory \( \Theta \) is more interpretive than alternatives if \( \Theta \) better reflects the ways that speakers of \( H \) represent the hypothesized semantic properties of \( H \)-expressions. As Heck (200x) notes, a cognitivist gloss also makes room for the idea that ‘true’ expresses “the very concept of truth that plays a central role in the semantic theory we tacitly know (p. 343).” Though as he shows, even if the mental language is the relevant metalanguage, very modest assumptions about the concept TRUE() still lead to paradox. Heck himself offers a different cognitivist gloss on (D), drawing on Higginbotham (198x) to suggest that one overt truth theory for \( H \) is more interpretive than another if the former better reflects how speakers of \( H \) use their semantic competence in communication. I have doubts about emphasizing communication; see section four. But this debate is intramural. And while we may not know how speakers represent truth conditions, evidence of various kinds can be relevant; see Evans (1982), Peacocke (1986) and Davies (1981, 1987), Pietroski et. al. (2009), Lidz et.al. (2011).
Note that if Lari had turned out to be a horse, as opposed to a Liar Sentence, one might have held that both sides of (1-T) are true. Correlatively, one might have thought that (1-T) is no more problematic than (17-T). And in my view, there is something deeply right about this last thought: (1-T) is not less true than (17-T). The difference is that we can approximate a proof that (1-T) is false, given that Lari is a Liar Sentence. And to repeat an earlier point, it isn’t paradoxical that (1-T) is false.

The right side of (1-T) is true, since Lari is not true. And since (1) is not true,

(1) Lari is not true.

the left side of (1-T) is false. But it’s OK for (1-T) to be false. This invented sentence lets us express a certain theoretical claim about the English sentence (1). And theoretical claims are often false, with no special exemptions for claims about human languages. We can invent languages in which the analogs of ‘true’ and ‘blue’ are of importantly different types, with the analog of (1) being ill-formed. But (1) is grammatical, and I see no reason for thinking that its grammatical structure is importantly different from that of (17).

(17) Bessie is not blue.

The word ‘true’ may well be polysemous, and perhaps homophonous. But this hardly shows that ‘true’ cannot have its ordinary semantic sense, whatever that is, in (1).

Perhaps (50) is true, and (1) is not TruthEvaluable in the relevant sense. That’s also OK.

(50) TruthEvaluable(‘Lari is not true.’) ⊨

True(‘Lari is not true.’) ≡ ~True(Linus)

One can say that (1) fails to be true in the way that horses and galaxies fail to be true—viz., by not being the sort of the thing that has a truth condition. If (1) doesn’t have a truth condition, then we need not and should not say that (1) is true if and only if Lari isn’t true. We can say that (51) is true, since both sides of the biconditional are false.

(51) True(‘Lari is not true.’) ≡ Precedes(Five, Two)

Just don’t replace the right side with a truth, as in (1-T), and try to maintain the biconditional.

Foster’s Challenge suggests that a meaning theory has to be governed by a naïve background “logic” that only permits derivation of theorems like (17-T) from simple schemata and substitutions. Liar Sentences suggest that a truth theory has to be sophisticated enough to not generate (1-T). Defense of (D) requires a plausible response to this apparent tension.

(D) For each human language H, there is a theory of truth

that is the core of correct theory of meaning for H.

This doesn’t show that (17-T) is false. But if a plausibly interpretive/cognized theory of truth has (17-T) as a theorem, then it is hard to see how the theory could fail to have (1-T) as a theorem. That is an indirect argument that no plausible theory of meaning for English will have (17-T) as a theorem. Similar remarks apply to (52-53).

(52) Lari is true. (52-T) True(‘Lari is true.’) ≡ True(Lari)

(53) Bessie is blue. (53-T) True(‘Bessie is blue.’) ≡ Blue(Bessie)

In my view, (52-T) is true, since both sides of this biconditional are false. And if Bessie is not blue, then (53-T) is also true. But if a plausibly interpretive/cognized theory of truth has (52-T) as a theorem, then it is hard to see how the theory could fail to (1-T) as a theorem; and if it has (53-T) as a theorem, then it is hard to see how the theory could fail to have (52-T) as a theorem. It would be tendentious to assume that the left side of (53-T) is false no matter what color Bessie is. But even if one can consistently maintain (53-T), it doesn’t follow that one should do so. The left side of (53-T) may be false because (53) doesn’t have a truth condition any more than (52) or (1) does. Correlatively, we need not say that (54) is true.
(54) ‘Bessie is blue.’ is true if and only if Bessie is blue. For we can say that (54) doesn’t have a truth condition any more than (1) does.

Let me stress this point. If (54) is a human language sentence—and so not confined to a written form governed by philosophical conventions—then it’s not at all clear that (54) is true, as opposed to a sentence that can be used to report the fact that (53) can be used to make a claim that is true if and only if the Bessie in question is blue in the relevant sense. It is often convenient to simplify, and say that (54) is true. But convenience is not an argument that instances of (55) are true, perhaps modulo context sensitivity; where ‘S’ is replaced by an English declarative and ‘S’ is replaced with that declarative quoted.

Even if ‘modulo context sensitivity’ can be cashed out plausibly for examples like (56),

(56) ‘I am hereby quoting him.’ is true if and only if I am hereby quoting him.

saying that instances of (55) are true is tantamount to adopting (D). If there are endlessly many true instances of some context-sensitive version of (55), as Davidson and many others suggest, then the obvious and perhaps only explanation for all these metalinguistic truths is that they follow from generalizations that constitute a truth theory for English. But the antecedent of this conditional is as much in question as (D) itself. Note that schema (55) has endlessly many instances like (57), whose quoted sentence can be used to report (some of) what Bert thinks.

(57) ‘Bert thinks that Ernie snores.’ is true if and only if Bert thinks Ernie snores.

If all these English biconditionals are true, then presumably, they follow from a finite number of basic principles that constitute a truth theory for English. But that’s no argument for the hypothesis that some theory of truth for English accommodates attitude reports.

Initially, it might seem that truth conditional semantics can be defended as the best explanation of the “fact” that instances of (57) are true, modulo context sensitivity. But this description of the explananda is tendentious. Liar T-Sentences like (58) highlight the point.

(58) ‘Lari is not true.’ is true if and only if Lari is not true. This instance of schema (55) might seem to be true if you don’t know what Lari is. But upon reflection, (58) seems to be a counterexample to (55). So one can’t insist that (54)

(54) ‘Bessie is blue.’ is true if and only if Bessie is blue.

is true on the grounds that all instances of (55) are true. On the contrary, one needs to take seriously the possibility that (54) is no more true than (1).

If H is a Tarski-style language, then H is related to the world in a certain (perhaps sophisticated) way. If H is a human language, then a human child has the capacities required to understand each expression of H, even if these capacities are quite unsophisticated. One can hypothesize that human biology threads this needle, letting members of our species acquire Tarski-style languages without special training. But prima facie, Liar Sentences tell against this hypothesis. And as we’ll see, “repairing” a simple theory—so that it is plausibly true—can be at odds with preserving its virtues as a hypothesized account of linguistic meaning.

2.4 Complicating Complicates

It is tempting to look for a theory that has theorems like (53-L) and (1-LT);

(53-LT) Legit(‘Bessie is blue’) ⊃ [True(‘Bessie is blue.’) ≡ Blue(Bessie)]

(1-LT) Legit(‘Lari is not true’) ⊃ [True(‘Lari is not true.’) ≡ ¬True(Lari)]

where both of these conditionals are true, as is the antecedent of (53-LT), but the antecedent of (1-LT) is false. Let’s not worry here about how to spell out the technical notion ‘Legit’—e.g., via fixed points (Kripke 1975), revisions (Gupta & Belnap 19xx), or whatever. For even if all the troublemakers can be filtered out, two related concerns remain if the goal is to defend (D).
(D) For each human language \( H \), there is a theory of truth that is the core of correct theory of meaning for \( H \).

First, it seems that any such filtering will yield consistency at the cost of explanation. If there are endlessly many English sentences such that the proffered theory does not assign interpretations to them, then we are left with the same kind of explananda that we started with. In thinking about human languages, the theoretical task is not merely to explain how a finite mind can understand boundlessly many expressions. One wants to know, among other things, how we humans understand the expressions that we do understand. If a theory of meaning accommodates endlessly many expressions like ‘chased a cow’ and ‘chased every brown cow’, but not ‘chased a cow today’, then the theory does not accommodate ‘chased a cow’ in the right way. And if a truth-theoretic account of meaning assigns an interpretation to (53) but not (15),

\[
(53) \text{Bessie is not blue.}
\]

(1) Lari is not true.

that is worrisome, since (1) is as meaningful/comprehensible as (53).

Moreover, and this is the second concern, theorems like (53-LT) and (1-LT) don’t seem to be interpretive. One can say that competent speakers recognize the antecedent of (53-LT) as true, that they discharge it to obtain (53-T), and that (1) is understood by analogical extension.

\[
(53-T) \text{True(‘Bessies is blue.’) = \text{Blue(Bessie)}}
\]

But if speakers can discharge the antecedent of (53-LT), then we need to revisit the concession that speakers tacitly deploy some truth theory \( \Theta \) that is governed by a very weak background logic that does not generate uninterpretive instances of (39a); where \( \Gamma \) is itself a theorem of \( \Theta \).

\[
(39a) \text{True(‘Ernie snores.’) = \text{Snores(Ernie) & \Gamma}}
\]

And in this regard, it’s important to remember that troublemakers like (1) cannot be filtered out in terms of grammatical properties; see Parsons (1974), Kripke (1975).

Adapting an example from Kripke, suppose that nine people wrote a sentence on a bit of paper that was put into an otherwise empty box called ‘Box’. The first four people wrote (3).

\[
(7) \ 2 + 2 = 5
\]

The next four wrote (20). The ninth person considered three options: (7), (20) and (58).

\[
(20) \ 2 + 2 = 4
\]

(58) Five bits of paper in Box carry inscriptions of a sentence that is not true. Once the last bit of paper is deposited, someone utters (58)—or another problematic sentence, say one that replaces ‘Five’ with ‘More than four’ or ‘An odd number of’. Unlike (7) and (20), (58) is context sensitive, if only because it is tensed. But it’s hard to see how any context sensitive element of (58) could track which sentence the ninth person chose to inscribe. Yet the utterance of (58) is false, true, or neither depending on this choice. So prima facie, (58) does not itself have a truth condition, not even if we fix values for any context sensitive elements.

If the ninth person inscribed (7) or (20), it might seem that (58) is false or true. But if that person inscribed (58), we see the problem with assuming that (58) has a truth condition. One can say that (58) is Legit in many but not all contexts. But then one needs to say how and when speakers discharge the antecedents of conditionals of the form \( \text{LEGIT}(S, C) \supset [\text{TRUE}(S) \equiv p] \).

Lycan (2012) recognizes that there is no syntactic characterization of the human language sentences that are Liarish troublemakers for truth-theoretic conceptions of meaning. But he offers a variant proposal: English is not a human language; rather, “English” is a hierarchy of sublanguages, each of which has its own proprietary truth predicate. The idea is that each speaker of English has a core competence that can be extended with a kind of metacapacity.
Think of the core competence as one that lets us generate (use, and comprehend) the expressions of a basic language $\lambda$, that is like the many dialects/idiolects we think of as versions of English, except for having a lexicon that contains no semantic vocabulary. The meta-capacity lets us generate the expressions of a closely related language $\lambda'$ that includes some semantic vocabulary, and in particular, a language-relative truth predicate $\text{T}_\lambda$ that is true of—or if you prefer, whose extension includes—all and only the sentences that are true-in-$\lambda$. For each expression of $\lambda$, there is a correspondingly structured and homophonous expression of $\lambda'$; and according to Lycan, a sentence of $\lambda$ is itself a sentence of $\lambda'$. So a sentence can be true-in-$\lambda$ and true-in-$\lambda'$. Sentences of $\lambda'$ that are not sentences of $\lambda$ cannot satisfy $\text{T}_\lambda$. But each sentence that is true-in-$\lambda'$ satisfies the language-relative truth predicate $\text{T}_{\lambda'}$ that appears in the language $\lambda''$, whose expressions are said to include those of $\lambda'$. And so on.

On Lycan’s view, “English speakers” have access to many truth predicates. But these predicates are not mere homophones, since a common conceptual thread runs through the hierarchy; although one cannot replace ‘true$\lambda$’, in a sentence of $\lambda'$, with ‘true$\lambda'$’. Lycan thus adopts a version of the idea that (1) is understood via analogy to sentences like (17).

(1) Lari is not true.

(17) Bessie is not blue.
The idea is that while (1) is not itself a sentence of any language in the “English hierarchy,” replacing ‘true’ with ‘true$\lambda$’ yields a sentence of $\lambda'$ that is true-in-$\lambda'$ but not a sentence of the basic language $\lambda$; though one cannot say this in $\lambda'$. Replacing ‘true’ in (15) with ‘true$\lambda$’ yields a sentence of $\lambda''$ that is true-in-$\lambda''$ but not a sentence of $\lambda$ or $\lambda'$; and so on. This may not preserve the idea that English (or any language in the hierarchy) has a truth predicate, at least not if such a predicate needs to have the meaning of ‘true’ (or be true of all and only true things). But advocates of truth-theoretic semantics might say that semantic words—perhaps including lexical items like ‘correct’, ‘honest’, ‘coherent’, and ‘heteronymous’—are special in this respect.

A deeper concern, in my view, is that Lycan’s proposal requires a single sentence to be a sentence of different languages. As discussed in section four, I think this is at odds with the idea that human languages are I-languages in Chomsky’s (1986) sense. Though as Lycan reminds us, Davidson held that a single utterance $u$ might be transcribed as /empedakliyz iypt/, classified by one interpreter as an utterance of the English sentence (59),

(59) Empedocles leaped.

and classified by another interpreter as an utterance of the German sentence (60).

(60) Empedocles liebt.

If $u$ can be true-in-English iff Empedocles leaped, and also true-in-German iff Empedocles loved, then $u$ can be true-in-English and true-in-German. So perhaps an utterance of (1) can be true-in-$\lambda'$, true-in-$\lambda''$, etc. We don’t view each utterance that might be transcribed as /bank/ as an act of uttering several words simultaneously. But in special cases, a person might perform a speech act that counts as “doing two things” with one utterance; and perhaps this point extends to cases of (unwittingly) uttering a sentence of both $\lambda'$ and $\lambda''$.

13 Only a few of which will ever be used in ordinary discourse. The suggestion is reminiscent of Davidson’s (1986) characterization of linguistic communication in terms of “passing” theories that speakers can deploy and share, in contexts, by adjusting “prior” theories that may not ever be used without adjustments of some kind. See Ludlow (2011) for related discussion.
Still, I don’t think a sentence of one I-language can be a sentence of another, at least not in the way that Lycan requires. It’s hard enough to make sense of the idea that (17) is a sentence of both a basic I-language and a lexically enriched I-language. (Does each speaker of English have an I-language that lacks the word ‘water’—and for any lexical items, an I-language that lacks them?) In any case, I don’t think (1) can be a sentence of two human I-languages that connect the articulation of ‘true’ to different semantic properties: humans acquire procedures that generate meaningful expressions, not mere syntactic structures; there is no “meaning-neutral” analog of (1) that a human I-language can generate. Nonetheless, Lycan has a serious proposal. So I’ll come back to it, after an initial discussion of context sensitivity and Davidson’s (1967b) claim that theories of truth for human languages will have theorems concerning utterances.

3. Relativizations or Actions
One might have hoped that the right account of context sensitivity will help in dealing with Liar Sentences. But as we’ll see, the two main strategies for accommodating sentences like (34)—

(37) I chased something.

relativizing the truth of sentences to contexts, or specifying truth conditions for sentential utterances—heighten the worry that no truth theory for a human language H is the core of a correct meaning theory for H. Focusing on truth favors the second strategy, according to which certain actions are truth-evaluable. Focusing on linguistic meaning favors the first strategy, according to which expressions are the real bearers of semantic properties, which are relativized. Thinking about the contrast reveals a tension in the idea that a single theory, concerning ordinary human beings, can be a good theoretical characterization of both truth and meaning.

3.1 Two Kinds of Theorems
As briefly discussed in chapter zero, we can extend a Tarskian language like PL (discussed in chapter two) in a Kaplanian way that yields models of sentences like (61) and (62).

(61) I chased it.

(62) I am here.

Add three indices—s, p, t—and a pointer ⇀ such that each pointer can be extended, with ′′′, to create another. Then expand each Tarskian sequence by adding a zeroth element that is itself an ordered triple of domain entities. Each expanded-sequence, or assignment of values to variables, is of the form: <e^′, e^′′, e^′′′, e, e^′, e^′′, ...>. For each assignment A: A(s)/A(p)/A(t) is the first/second/third element of A’s zeroth element; and each pointer has the value of the corresponding variable—A(微软) = A(x), A(微软′) = A(x′), etc. This allows for object language sentences like (61a) and (62a), along with metalanguage theorems like (61b) and (62b).

(61a) CHASED(s,微软)

(61b) Satisfies[A, CHASED(s,微软)] ≡ A(s) chased A(x)

(62a) LOCATION(s, p)

(59b) Satisfies[A, LOCATION(s, p)] ≡ A(s) is located at A(p)

This leaves room for further analysis. But whatever the details, (58a) and (59a) are not true or false, since these sentences are satisfied by some but not all assignments. One can, however, say that contexts license assignments. Suppose that whatever contexts are, each context

14 See chapter three. For example, (61a) might be decomposed as \( \exists X^*(\text{BEFORE}(X^*, t) \& \text{CHASE}(X^*, s,微软)) \), which is satisfied by \( A \) if and only if: some \( A^* \) such that \( A^* \sim X^* \) \( A \) is such that \( A^*(x^*) \) occurs before \( A^*(t) \), and \( A^*(x^*) \) is a chase by \( A^*(s) \) of \( A^*(x) \); or abbreviating, something “was” a chase by \( A(s) \) of \( A(x) \).
determines exactly one speaker, place, and time, along with some (perhaps null) n-tuple of deictic actions. Then one can say that context \( c \) licenses assignment \( \mathcal{A} \) if and only if: \( \mathcal{A}(s/p/t) \) is the speaker/place/time of \( c \); and if \( \Delta \) is the \( i \)th deictic action determined by \( c \), then the \( i \)th element of \( \mathcal{A} \) is the domain entity demonstrated via \( \Delta \). If \( c \) determines a deictic action in which nothing is demonstrated, then \( c \) does not licence any assignment. But bracketing this possibility, one might adopt the context-relative notion of truth characterized in (63).

\[
(63) \quad \text{True}(S, c) \equiv \text{for each assignment } \mathcal{A} \text{ that } c \text{ licenses, } \text{Satisfies}(\mathcal{A}, S)
\]

If each conversational situation determines a unique context, then modulo vagueness regarding the relevant speaker/place/time, one can say that a sentence \( S \) is true in that situation if and only if \( S \) is satisfied by each assignment that is licensed by the determined context. This allows for various conceptions of how (knowledge of) meaning is related to (knowledge of) pragmatic principles concerning the use of expressions in conversational situations.

This shows, among other things, that mere inclusion of an indexical or deictic expression does not preclude a sentence from having a context-relative truth condition. Given a Kaplanian extension of a Tarskian language, T-theorems still concern the truth conditions of sentences. And one can hypothesize that human languages are variants of some such extension of PL. But it doesn’t follow that (61) and (62) are true or false, relative to contexts or in situations, since human language sentences may not be truth-evaluable at all. Providing models for (61) and (62) does not show that these sentences have truth conditions relative to contexts. Even ignoring the possibility of failed acts of demonstration, issues concerning chases and event variables remain; see chapter three. Moreover, even if objections specific to (58) and (59) can be addressed, our question is whether there are meaning-specifying theories of truth for human languages, not whether sentences like (61) and (62) themselves favor a negative answer.

It is also worth remembering that according to Davidson (1967b), a truth theory for a human language will have theorems concerning utterances; cp. Burge (197x), Higginbotham (1985), Larson and Segal (1995), Lepore and Ludwig (200x). Regarding (61), one might suggest generalization (61c); where ‘\( \text{True}[u, S] \)’ means that \( u \) is a true utterance of sentence \( S \),

\[
(61c) \quad \text{True}[u, \text{CHASED}(s, \infty)] \equiv \text{the author of } u \text{ chased the (first and perhaps only) thing demonstrated in } u \text{ and an utterance of (61) can be an utterance of } \text{CHASED}(s, \infty). \text{The idea is that an utterance is an action that can be described as a spatiotemporally located articulation, by a speaker in a place at a time, of a sentence that has a certain logical form. For many purposes, the D(avidson)-strategy and the K(aplan)-strategy are interchangeable. But here, I want to stress that they reflect different ways of thinking about how human languages and their users are related to truth and contexts.}
\]

On the K-strategy, a theory of truth for a language is a theory of a language. And for these purposes, one can even take languages to be sets of expressions rather than generative procedures. The important point here is that on the K-strategy, theories of truth are theories of linguistic objects, independent of how they are used. Uttering a sentence, in a certain situation, may determine which assignments are licensed on that occasion of use; and in this sense, truth depends on use. But the context-sensitive truth conditions of a sentence \( S \) are specified without reference to actions of using \( S \), except perhaps with regard to deixis.

By contrast, on the D-strategy, a theory of truth for a language is a theory of certain spatiotemporally located actions of using the language. The theory ascribes truth conditions to utterances of generable sentences. This approach may locate truth in a more plausible place. But
it is no small task to formulate axioms from which biconditionals concerning utterances follow, given a weak logic that avoids unwanted/uninterpretive theorems.

Accommodating quantificational constructions like (37) is hard enough,

\[ (34) \text{I chased something.} \]
even with logical forms like (37a), given the need for sequence relativization as shown in (37b).

\[ (37a) \exists x (\text{chased}(s, x)) \]

\[ (37b) \text{Satisfies}\{A, \exists x (\text{chased}(s, x))\} \equiv \]

some \( A^* \) such that \( A^* \models x \exists A \text{ is such that } A(s) \text{ chased } A^*(x) \)

So, one might prefer a Kaplanian theory that delivers theorems like (37b) rather than (37c),

\[ (37c) \text{True}[u, \exists x (\text{chased}(s, x))] \equiv \]

some \( A^* \) such that \( A^* \models x \exists A \text{ is such that the author of } u \text{ chased } A^*(x) \)

here ignoring event variables; cp. note 14. The added complexity of (34c) has to be reflected, one way or another, in axioms from which such theorems are allegedly derived. Specifying sentence meanings in terms of truth-evaluable utterances also seems to confuse linguistic competence with language use, even if we bracket utterances that are neither true nor false.

Following Wittgenstein and others, one can hypothesize that linguistic competence is competence regarding the use of expressions, as opposed to knowledge of expressions whose properties constrain how they can be used. But recognizing implications, for example, does not seem to be a matter of generalizing across utterances. Moreover, many—perhaps most—uses of human linguistic expressions are intrapersonal. And even if truth-evaluable utterances play a central role communication, truth-evaluable judgments may not play an analogous role in cognition. One can speculate that using expressions to entertain thoughts (“just thinking”) is derivative on using expressions to endorse thoughts. But prima facie, it’s the other way around.\(^{15}\)

It can be tempting to say that languages are for making assertions. Though we shouldn’t assume that human languages are “for” any particular use, especially if they turn out to be I-languages in Chomsky’s (1986) sense: biologically implemented expression-generating procedures that pair articulations with meanings; see section four. One can hypothesize that human I-languages pair articulations with mental analogs of (37c), as opposed to (37b), despite intrapersonal uses of I-languages. But another possibility is that I-languages are used for many purposes, including the generation and construal of certain utterances, without speakers encoding any theory of meaning that employs a truth predicate that applies to utterances. In short, it may not be independently plausible that I-language expression meanings are specified by axioms concerning utterances (or any truth-evaluable uses) of the expressions in question.

### 3.2 Evaluating Actions

On the other hand, the D-strategy can seem reasonable if the goal is to specify theories of truth for human languages. With regard to such languages, it may well be that truth is more plausibly ascribed to certain actions of using sentences than to sentence-context pairs.

From a Strawsonian perspective, (58) vividly exemplifies a general point.

\[ (58) \text{Five bits of paper in Box carry inscriptions of a sentence that is not true.} \]

A declarative sentence S can fail to be true or false, even given values for any variables in S. One can say that sentences like (64) carry presuppositions—e.g., that the name denotes something—

\[ (64) \text{Vulcan is bigger than that.} \]

\(^{15}\) Perhaps all mental uses of language are somehow derivative on episodes of interpersonal communication. But I don’t of any good reasons, as opposed to “private language arguments,” for thinking that this is so.
and that derivable specifications of truth conditions are correspondingly conditionalized. But then the allegedly meaning-specifying theorems (and axioms) are correspondingly complicated. Moreover, while some presuppositions seem to be grammatically constrained, it is hard to account for even these constraints on truth-evaluability in terms of properties of the expressions themselves, as opposed to general knowledge of how expressions can be used to convey information; see Schlenker (2009). In any case, it’s hard to see what presupposition of sentence (58) is violated if one bit of paper in Box contains an inscription of (58).

Prima facie, some constraints on truth-evaluability are constraints on actions of using sentences to make assertions or judgments in particular contexts. I think that many observations regarding vagueness also point in this direction. In a context where it is clear what counts as being green—or least as clear as it can be for normal humans—there can still be vague cases. Following Sainsbury (1996) and many others, I conclude that ‘green’ does not have an extension, not even relative to contexts. No set is such that for each thing x, x is an element of that set if and only if x is green. And I see no independent reason for thinking that the meaning of ‘green’ determines a function from contexts to extensions, as if ‘green’ contains a covert variable whose values correspond to ways of replacing a vague predicate with a precise one.

Supervaluation can, of course, be useful for certain purposes; see, e.g., Fine (19xx). In particular, when humans try to make truth-evaluable claims by using expressions that do not have extensions, vague cases can present complications or even paradoxes. In some theoretical contexts, it can be helpful to regiment such usage and bracket certain questions about meaning. This may improve our understanding of how the practice of making an initial thought more precise is itself sensitive to context. But as Wright (19xx) and others have discussed, many uses of human language predicates are intentionally tolerant. In using ‘green’, we often want to avoid making an assertion whose truth or falsity could depend on a distinction that ordinary humans cannot make.16 In many contexts, a person who thinks that Kermit is green can be described as someone who endorses the truth-evaluable proposition that Kermit is green. But such description does not commit us to saying that ‘green’ is true of Kermit.

These considerations suggest that the true/false distinction, like some other right/wrong distinctions, targets actions that meet the conditions for being evaluable in a certain way. We can idealize, say that human language sentences are true or false relative to contexts, and then speak of propositions as abstracta that are unrelativizedly true or false. Given episodes of thinking, we can speak of thoughts thought. Given episodes of judging, we can speak of propositions judged true; see, e.g., Cartwright (1962). But it doesn’t follow that when actions of assertion/judgment are true or false, they are truth-evaluable by virtue of having propositional contents, much less that theories of meaning should specify such contents. If truth lies downstream of meaning—in that acts of using sentences are candidates for being true or false, subject to further review—then perhaps theories of meaning should not deploy the predicate ‘True(_ )’, even if humans enjoy a mental predicate TRUE(_) that applies to certain assertions/judgments. That leaves the hard task of describing the (presumably context sensitive) conditions on being true or false, and explaining

---

16 Similar points apply to the practice of making laws by using words like ‘painful’, ‘lethal’, or ‘intentional’ in official contexts. See Graff (2000) for important related discussion. Note that inscribing ‘{x: x is green}’ does not establish that the inscription has a denotation; cp. ‘{x: ¬(x ∈ x)}’. See Pietroski (2005a) for related discussion in the context of Boolos’ (1998) conception of plural quantification. Perhaps when doing logic, uses of concepts are subject to normative constraints according to which GREEN(_) is to be used as if it has an extension, even if don’t know which extension this is; see Williamson (1994). But even given some version of “epistemicism” for ideal(ized uses of) concepts, it doesn’t follow that the human language word ‘green’ has an extension.
why these are conditions on truth/falsity; but see Glanzberg (2004, forthcoming) for a promising start.

One can agree with Davidson that a truth theory for a human language $H$, if such there be, will have theorems that specify the truth conditions of certain spatiotemporally located actions of using expressions of $H$ (rather than use-independent properties of those expressions). But one might also think that a good theory of meaning for $H$ needs to have theorems that specify use-independent properties of the boundlessly many expressions of $H$; where these features of $H$-expressions constrain and partly explain how the expressions can and cannot be used. For purposes of studying specific constructions like (68),

(68) Two chipmunks chased each other around the park for an hour.

it may not matter which way advocates of (D) accommodate indexicals and demonstratives.\(^{17}\)

(D1) for each human language $H$, there is a correct theory of truth for $H$.

But in evaluating (D), one cannot ignore questions about how meaning and truth are related,

(D) For each human language $H$, there is a theory of truth that is the core of correct theory of meaning for $H$, or how they are jointly related to sentences, utterances, and contexts.

Recall that Russell’s (190x) notion of a proposition led Frege (190x) to ask if concrete particulars like Mont Blanc could be constituents of truth-gradable entities. Likewise, Strawson rejected the idea that meanings are referents. From his perspective, Russell mischaracterized the sense in which judgments about mountains are spatiotemporally located: the judgeable contents do not have constituents that can be climbed. In this respect, judgments about mountains and their snowfields are as abstract as judgments about numbers and their successors; contents are individuated finely, like ways of thinking about things (see chapter one). But an episode of judging that every number has a successor is still spatiotemporally located, and evaluable as true, even though the constituent concepts let us think about abstracta in certain ways.

Like ‘thought’, ‘expression’ exhibits symbol/episode polysemy. An invented formula can be an expression of PL. Writing a note of condolence can be an expression of sympathy. Human sentences are expressions, as are many actions of using such sentences. This can be confusing, since acts of endorsing or conveying truth-gradable thoughts can be evaluated in many ways. A sentence can be used to make judgments/assertions that are rash, rude, rebellious, true, etc. But are the sentences themselves rash, rude, rebellious, or true relative to contexts? In addressing such questions, we need to think about what human languages and their expressions are.\(^{18}\)

One can hypothesize that humans have both a capacity for acquiring I-languages, and an ancillary system that specifies—or in some other sense determines—truth conditions for certain uses of I-language expressions that already pair articulations with meanings. And one can try to specify these meanings, which human languages pair with articulations, in truth-theoretic terms via some version of the Kaplan-strategy that allows for certain dimensions of context sensitivity. This is an interesting hypothesis, especially given an explicit proposal about the requisite very weak background logic deployed to generate theorems like (21-T); see section two.

(21-T) True(‘Ernie snores.’) = Snores(Ernie)

---

\(^{17}\) Many of the virtues and vices translate. See Glanzberg (201x) for related discussion in the context of comparing truth-theoretic and model-theoretic conceptions of semantics; cp. Lepore (1983).

\(^{18}\) It may be no accident that when Davidson (1986) addressed the question of what truth theories are theories of, he concluded that they are not theories of languages in any traditional sense; see Pietroski (1994) for discussion.
But then one can’t deal with Liar Sentences by saying that truth and falsity are really properties of spatiotemporally located assertions/judgments, and that truth theories for human languages assign semantic properties to utterances rather than expressions relativized to contexts.

3.3 Contexts, Kettles, and ‘That’-Clauses
In chapter zero, I reviewed some reasons for thinking that the truth or falsity of a thought expressed with a sentence like (69) or (70) can depend on the context of use in subtle ways.

(69) The kettle is black.
(70) Sam denied that Hesperus is Phosphorus.

I want to recapitulate this point before getting back to Liar Sentences in section four. While this may seem like a digression, my intention is to connect the details of this chapter with the larger theme of the book.

In my view, the meaning of a human linguistic expression $\Sigma - \mu(\Sigma)$ for short—is an instruction that determines the form of a concept assembled by executing this instruction. A concept that is assembled by executing $\mu(\Sigma)$ in a given context can also be a concept that is expressed by using the pronunciation of $\Sigma - \mu(\Sigma)$ for short—in that context. But executing $\mu(\Sigma)$ on other occasions may lead to different results if at least one of $\Sigma$’s lexical constituents has more than one fetchable concept at its address. Lexical meanings constrain without determining concept choice. In this sense, the concept associated with $\Sigma$ can vary across contexts, even if no index in $\Sigma$ tracks this dimension of variation. Likewise, the sound associated with $\pi(\Sigma)$ can vary across contexts, even if no index in $\Sigma$ tracks the difference between shouting quickly at the top of one’s vocal range or speaking slowly and softly at the bottom of one’s vocal range.

If a word like ‘France’ (or ‘Venice’) can be used to access a concept of a certain geographical region, or a concept of a certain political unit, then we can make sense of why sentence (71) is weird in a way that (72) is not.

(71) France is a hexagonal republic.
(72) France is hexagonal, and France is a republic.

For each use of ‘France’ in (72), there is a way of executing the instruction that leads to a reasonable thought; but not so for the single use in (71). This also helps explain why inferring from (72) to (71) feels wrong, or at least risky. A thought constructed in accord with (72) can be true; but a thought constructed in accord with (71) is likely to be false, at least in many contexts. But if ‘France’ does not determine which concept is to be fetched via this proper noun, then (73)

(73) The author of Conjoining Meanings likes France.

expresses neither a thought about certain terrain nor a thought about a certain nation. More generally, I claim, there is no thought that (73) expresses and no truth condition that (73) has.

In a particular context, there may be reasons for using ‘France’ to access the political concept. But even in such a context, I don’t think the meaning and context jointly determine the concept accessed, unless one builds the accessing into the context. And I see no reason for thinking that ‘France’ somehow includes an index whose character somehow tracks which concept one ought to access in a given context. In many contexts, there may be a fact of the matter concerning which concept one ought to access for purposes of reasonable conversation; and the meaning of ‘France’ certainly constrains the space of possible (human) answers to the question of how this instruction should be executed in a given context. But the meaning/instruction is not itself a mapping from contexts to answers.

Meaning constrains the use of expressions while leaving room for various kinds of cognitive freedom. Travis (199x) makes this point with examples like (69).

(69) The kettle is black.
And while he does not describe lexical meanings as instructions for how to access concepts, I think an important moral of his discussion can be put in these terms: when a hearer encounters a sentence that contains a certain lexical item L, she often has a choice regarding which concept to access with L; and often, her choice will be influenced by the rest of the sentence and the larger context. This suggests that we need to distinguish the compositionally determined meaning of a sentence from the use of that meaning in constructing a potentially truth-evaluable thought.

Chomsky (1977, 2000b) makes similar points; see Pietroski (2003, 2005b). Carston (2002) speaks of “ad hoc concepts,” in order to stress that in speech or comprehension, one may restrict or relax available concepts for specific temporary purposes; see also Wilson and Sperber (200x). This is fully compatible with my conception of meaning. Lexical addresses do not harden and refuse to admit new residents. On the contrary, if children use some concepts to introduce others in the course of lexicalization, then we shouldn’t be surprised if the use of a lexical item L in communication often involves using the concepts at L’s address to introduce yet another concept whose utility may be fleeting. Indeed, the processes that support lexicalization may be continuous with those that support communication; cp. Davidson (1986), Ludlow (201x).

One wants to know how much freedom we have to introduce concepts in ad hoc fashion. One also wants to know what concepts are. But our ignorance is not a reason for skipping over concepts and saying that lexical items map contexts to extensions. That said, let me stress that I take meanings to be instructions, not their executions. A hearer might recognize the meaning of an uttered expression \( \Sigma \), use \( \Sigma \) to build a concept, yet also recognize that the speaker would use \( \Sigma \) to build a different concept of the specified form (with constituents that are somehow more restricted/relaxed than those currently available to the hearer). Sometimes, understanding a speaker’s words triggers appreciation of the fact that the speaker uses those words to access concepts that are subtly but perhaps importantly different than any concepts one currently has access to. Reading philosophy seems like a paradigm case.

It is a fantasy—arguably a horrific one—that all speakers of English attach the same concepts (or the same mappings from contexts to extensions) to English expressions. Even ignoring the differences across I-languages that count as English, acquiring such an I-language leaves room for variation in the concepts that individual speakers connect with particular lexical items; and it leaves room for choice with respect to which concepts one accesses on a given occasion of using some lexical items to access some concepts. Pace Dummett (1986), this does not imply that successful communication is accidental in any worrying sense. As members of the same species, humans share many cognitive dispositions. The fact that other primates do not acquire pronounceable lexical items, while children acquire them quickly and with apparent ease, suggests specialized circuitry that every human enjoys. (If we meet intelligent aliens who are not biologically based, talking with them may require more luck.)

Speakers of English connect lexical items with concepts in ways that are good enough for successful human communication, at least often. But success differs from accuracy; and for most practical purposes, I prefer success. The concept a speaker uses, on a given occasion of using ‘black’ to talk about a kettle, probably doesn’t have the same extension as the concept that any hearer uses. That doesn’t preclude (tacit) agreement that the kettle in question counts as black for the purposes at hand. Correlatively, while human linguistic communication is often successful in many ways, this does not justify the claim that ‘black’ maps contexts to extensions.

One can hypothesize that lexical items have (or are) indices that have Kaplanian characters. But even if the lexical item ‘black’ is always used with an accompanying index \( i \), yielding a complex homophonic expression of the form [black-\( i \)], that leaves the task of
specifying the meanings of the lexical item and the index. This task is difficult, even if we assume that [black-i] is of the Fregean type <e, t> and impose no constraints on lexical types. This allows the index to be of any Fregean type <α>, with the lexical item being of the corresponding type <α, <e, t>>. But however one encodes the alleged dimension of contexts, allegedly tracked by the alleged index, one still has to encode the context-invariant contribution of the lexical item. Moreover, even if this unconstrained homework problem is solvable, there are constraints on lexical types, even if we don’t yet have a good account of them. Positing indices of any unusual type—say, other than the one(s) posited for ‘I’, ‘here’, and ‘now’—raises the question of whether human I-languages permit such indices; see Chierchia (198x).

One can speculate that all open-class lexical items combine with unpronounced indices of type <e>. But human I-languages may not permit lexical items that express the requisite relations between entities like kettles and the (presumably abstract) entities that allegedly reflect dimensions of contexts (e.g., the alleged blackness-dimension) in the way that ‘I’ reflects the speaker-dimension of contexts. Even if theorists can specify such relations, and stipulatively assign them to invented models of lexical items, it hardly follows that lexical items specify such relations. I can’t prove that children don’t understand ‘black’ in terms of a dyadic relation that a kettle can bear to some entity that effectively selects one of the potential extensions for ‘black’. But prima facie, the context dependence of truth goes beyond the context sensitivity of meaning.

As noted in chapter zero, reports made with sentences like (70)

(70) Sam denied that Hesperus is Phosphorus.

present analogous but more severe difficulties for thesis (D).

(D) For each human language H, there is a theory of truth that is the core of correct theory of meaning for H.

This is so, even if the “opacity” of ‘that’-clauses can be accommodated. Initially, one might think that the context sensitivity of indirect speech reports (as opposed to direct quotation) is due to a special composition rule—a rule that has the effect of introducing a covert index—for phrases in which a verb combines with a ‘that’-clause. It can also be tempting to say that combining an adjective with a noun, as in ‘black kettle’, has the effect of introducing an index; though even if this helps with (74), (75) also illustrates the context sensitivity of ‘black’.

(74) The black kettle is on the table.

(75) The kettle on the table is black.

And while indirect speech reports may be special cases, in several respects, the phenomenon of opacity is not a license to posit a composition rule that is sensitive to context. More generally, the context sensitivity of truth is not a license to assume that meanings are compositional only in a weak sense that is compatible with (D) given examples like (70) and (75).

We can invent languages that allow for sentences like (76),

(76) Sam denighed glonk that Hesperus is Phosphorus.

in which ‘denied’ is a ditransitive verb whose indirect object ‘glongk’ is an index that tracks which substitutions within ‘that Hesperus is Phosphorus’ are permissible in the relevant context. We can stipulate that ‘denigh’ is true of <ε, β, γ> relative iff ε is an event of denial whose “content” is β relativized to γ; where β is the sort of thing denoted by a ‘that’-clause, and γ is a potential value of the index ‘glongk’. Then one can say that (76) is true relative to a context c iff some denial is such that it was done by Sam prior to c(‘now’), and its content is the denotation of ‘that Hesperus is Phosphorus’ as relativized to c(‘glongk’y). One can hypothesize that (70) has the
same truth condition. This is to posit a sophisticated and hard-working composition rule for phrases in which a verb combines with a ‘that’-clause, raising the question of whether human linguistic meanings can combine in this way—and if they can, why this kind of sophisticated and hard-working composition rule is not more often attested in human languages.

For these purposes, let the meaning of ‘that Hesperus is Phosphorus’ be whatever you like. I take ‘that’-clauses to be instructions for building mental predicates of a special sort: predicates that apply to certain concepts that have been prepared for declarative (as opposed to interrogative) use; see Lohndal and Pietroski (2011). As discussed in chapter six, I think the meaning of ‘deny that Hesperus is Phosphorus’ is an instruction for how to build a concept of the form \( \text{DENY}(\_)^\exists[\text{CONTENT}(\_ , \_ )^\Phi(\_ )] \); where \( \text{CONTENT}(\_ , \_ ) \) is a thematic concept akin to \( \text{PATIENT}(\_ , \_ ) \). But if you like, let ‘that Hesperus is Phosphorus’ denote some fine-grained proposition, perhaps the Fregean sense of an idealized version of ‘Hesperus is Phosphorus’.

The point here is not merely that the meaning of (70) differs from that of (77).

(77) Sam denied that Hesperus is Hesperus.

The point is that the meaning of (70) does not determine a mapping from contexts to truth values. In light of work by Davidson (1967a) and others, we should not reject thesis (D) just because certain substitutions do not preserve truth. Obviously, (78) does not imply (79).

(78) Slim is so called because he is thin; and Slim is Jim.

(79) Jim is so called because he is thin.

Likewise, (80) does not imply (81), and (82) does not imply (83).

(80) Slim is Jim; and Kim said, ‘Slim is thin’.

(81) Slim is Jim; and Kim said, ‘Jim is thin’.

(82) Slim is Jim, and Kim said so.

(83) Jim is Jim, and Kim said so.

But these nonimplications just highlight the fact that human languages allow for logophors that are in many respects like markers for direct quotation; see Quine (195x), Cartwright (1962), Forbes (199x), Pietroski (1996). So even substitution of synonyms is not always valid. We can say that ‘lawyer’ does not rhyme with ‘attorney’. And while quotation raises puzzles, I don’t think that the quotational aspects of (78-83) present especially deep difficulties for (D). But this isn’t to grant that (77-84) have compositionally determined truth conditions.

(84) Kim said that Slim is thin.

Suppose that Kim uttered (85) at some time \( t \), using ‘James’ as a name for Slim/Jim.

(85) James is gaunt.

Then an utterance of (84) might be true. If the intended audience for (84) only knows the man in question as Slim, and Kim used ‘gaunt’ only after checking a thesaurus, then (84) may be a good tool for reporting Kim’s utterance of (85). But in a different situation, an utterance of (84) might be false. If Kim’s choice of ‘gaunt’ as opposed to ‘thin’ was medically important, and everyone

---

19 Cp. Schiffer (1992), Ludlow (199x). Citations to Higginbotham, Segal, Larson/Ludlow on appeals to similarity.

20 See Cappelen and Lepore (2005). Sentence (89) can be pronounced as ‘Kim said, quote, Slim is thin’. And many human languages have a distinctive complementizer for direct speech reports; see Munro (198x) for interesting discussion. Of course, complementizers differ from demonstratives in many ways. In particular, a complementizer has to target the sentence it introduces, as illustrated with (92), which does not have the meaning of (92a); see Segal and Speas (1986), Seymour (1994).

(92) Slim is thin. His mother said that he needs to gain weight.

(92a) #Slim is thin. His mother said that. He needs to gain weight.

But the spirit of Davidson’s account does not require that complementizers be demonstratives; see Pietroski (1996).
thinks that Slim is Jim’s overweight cousin, then an utterance of (84) would not correctly report Kim’s speech act. Likewise, if Kim uttered (86), some utterances of (87) might be true,

(86) Slim is thin.
(87) Kim said that James is thin.

while other utterances of (87) are false.

Davidson himself stressed such points, and so appealed to a notion of “samesaying” that dovetailed with his inclination to ascribe truth to certain actions rather than sentences relativized to contexts. On his view, an utterance of (84) is true iff the embedded utterance of (86) “samesays with” some action of saying done by Kim. As noted in chapter zero, one can recode this idea by saying that ‘said’ expresses a relation that sayers bear to sentences relativized to contexts. But there remains the question of what it is to “say” a sentence relative to a context. So while I see how to invent languages that allow for opacity, that doesn’t speak to real challenge that sentences like (87) and (75) present for the idea that such sentences have truth conditions.

(75) The kettle on the table is black.

3.4 Contexts and Commonsense

One can try to argue that sentences of a human language contain the indices required by (D),

(D) for each human language H, some theory of truth

is the core of a correct theory of meaning for H

by offering independent arguments for (D) and the further claim that the most plausible version of (D) is one that preserves the following idea: if an utterance of some sentence S is true in one context, but an utterance of S is false another context, then S contains an index that has different values relative to the different contexts; see Stanley (2000, 2002). But then the argument for the requisite indices relies on defense of a thesis bolder than (D). And note that even if sentences of a human language include enough indices to track the ways in which truth can depend on context, it doesn’t follow that these indices can take the requisite values, or that human languages can support the requisite modes of composition. And given (C),

(C) each human language is an I-language in Chomsky’s sense

limitations on lexical meanings and semantic composition may severely constrain the possible values of any posited indices.

As noted in chapter one, and further discussed in chapter six, it’s hard enough to figure out which conjunction and closure operations are employed in human semantic composition. But these biologically implemented operations are surely limited in ways that impose constraints on the space of possible meanings for expressions (atomic and complex) of human I-languages. So to argue for covert indices by arguing for (D), one needs to argue for (B)

(B) each human language H is an I-language such that

some theory of truth is the core of a correct theory of meaning for H

without assuming that human I-languages generate expressions with the requisite indices. If (B) was independently plausible, modulo a few special cases, it might be reasonable to posit some otherwise unmotivated indices. But if every sentence of a human language seems like a counterexample to (B), unless one idealizes away from the problematic aspects of that sentence, then I don’t see how to use (D) as a bootstrap for justifying appeal to covert indices that help respond to some—but by no means all—of the objections to (B).21

21 At this point, debate can degenerate into trading hunches about whether to describe particular sentences as counterexamples to (D), or as invitations to pursue Stanley’s strategy, treating every example as a potential research program. So I don’t think it will be useful to rehearse more examples. But there remains the challenge of defending
Some advocates of (D) explicitly reject appeal to covert indices, absent independent motivations for such appeal, and conclude that a single utterance of a sentence can be an act of expressing more than one proposition; see Cappelen and Lepore (2005). If ‘black’ contains no index, one can hypothesize that each utterance of (69)

(69) The kettle is black.

has a “minimal” (or primary) truth condition that can be specified in some preferred theoretical context, but that an utterance of (69) can also have another truth condition that is more specific to the context of that utterance. But if uses of ‘black’ exhibit various truth-theoretic properties, why think that the lexical item also exhibits a truth-theoretic property that is constant across such uses? Why posit the alleged minimal/primary truth conditions?

In principle, one might propose algorithms for computing “secondary propositions” from “primary propositions” and representations of certain aspects of situations that speakers recognize as relevant, despite the absence of corresponding linguistic indices. But as Cappelen and Lepore admit, examples like (69) leave little hope of specifying plausible algorithms of this sort. Cappelen and Lepore defend their proposal as the best option for advocates of (D), absent evidence that words like ‘black’ include indices that track the relevant aspects of contexts. But this is to admit boundlessly many cases of a sentential utterance having a truth condition that is not characterized by a theory of truth/meaning for the language in question. Prima facie, this threatens the conception of meaning characterized by (D).22 And especially given thesis (C), Stanley might argue that his Kaplan-esque strategy for defending (D) by appeal to covert indices—and taking truth to be a property of sentences relativized to contexts—is better than taking truth to be a property of utterances and saying that an utterance can have more than one truth condition. But even if one version of (D) is more plausible than others, that is not yet an argument for (D) or a response to the apparent counterexamples.

We can speak of first meanings, second meanings, and so on; see Davidson (1986). Or drawing an analogy to Marr’s (1982) description of visual processing, one might say that executing what I call a meaning yields a “primal sketch” of a thought, which can be an input to further processing that yields a multi-dimensional representation—perhaps one that includes a Schein-style logical form of the sort discussed at the end of chapter three. But prima facie, the meanings that human I-languages connect with pronunciations do not track all the ways in which the truth of falsity of assertions (made in part by using the meanings) can depend on the context.

It is also worth remembering that examples like (69) and (88)

(69) The kettle is black.
(88) Al chased Theo.

were supposed to be cases that motivate (D), despite all the examples suggesting that the distinction between metaphorical and “literal” commonsense talk is vague and context sensitive. So if (69) and (88) already illustrate reasons for rejecting (D), that is significant, since ordinary speech is often much less friendly to (D). For example, consider the lines below, from Robert Mellin’s lyric for ‘My One and Only Love’.

---

(D) without positing powerful apparatus that is not independently motivated. And this challenge is heightened given an alternative conception of semantic composition that is more constrained.

22 It seems like a Pyrrhic victory for Davidsonians if (D) is preserved by saying that (i) a single utterance of a sentence can have more than one truth condition, and (ii) the truth condition that matters for interpretation of the utterance—a human action—differs from the truth condition determined by the lexical axioms and composition principles, even given contextually determined values for any variables in the expression.
(89a) The very thought of you makes my heart sing like an April breeze on the wings of spring.
(89b) And you appear in all your splendor.
(90a) You fill my eager heart with such desire. Every kiss you give sets my soul on fire.
(90b) I give myself in sweet surrender.

Let’s agree to set (89a) and (90a) aside as metaphorical. Hearts don’t sing; neither are they eager or filled with desire. Spring has no wings. There are no souls; and even if there were, they couldn’t be set on fire. So nobody owes a theory according to (89a) and (90a) are true in the relevant context. By contrast, while (89b) and (90b) are stylized, I don’t think they are metaphorical, at least not in the ways that (84a) and (85a) are. One can extend the notion of metaphor so that it covers just about all ordinary discourse, and then try to characterize the apparent distinction between the (a)-cases and (b)-cases in some other way. But if advocates of (D) end up saying that almost all sentences are false in most contexts, and ordinary intuitions of “correctness” are correspondingly unreliable indicators of truth, then we need to hear much more about the motivations for (D). So I assume that at least prima facie, advocates of (D) should count (89b) and (90b) as potential truths.

In which case, a person α—or α’s appearance, or an event of α appearing—can be in all α’s splendor, and so α has some splendor that x can be in; cp. talk of chases, sunrises, cites, and skies. Likewise, an event of β giving β (perhaps to α) can apparently be in sweet surrender. Regardless of whether these are examples of polysemy or homophony, a theory of truth for English would need to specify satisfiers for ‘in’ accordingly. In my view, disquotational specification is adequate. But in any case, (D) seems to require a domain that includes entities like appearances/givings that are somehow “in” things (e.g. splendor/surrender) in ways that are harder to comprehend than the corresponding expressions. And even if one posits the requisite “things,” there remains the question of how ‘in’ could be true of the requisite ordered pairs.

One can say that (89b) and (90b) are quasi-metaphorical special cases. But even if no single family of examples shows that (D) is false, each family can add to the overall argument that meaning is not so tightly connected to truth. In my view, the connection is indirect—via the use of assembled concepts that are used in context dependent ways—and in many cases, quite tenuous. I think Liar Sentences make it possible to see this point through a narrow but sharply focused lens. The argument that (1) has no truth condition approximates a proof.

(1) Lari is not true.
This doesn’t establish that (53) has no truth condition.
(53) Bessie is blue.

---

23 Cp. Lakoff (19xx). Perhaps “in the hush of night while you’re in my arms” is doubly metaphorical: night has no hush to be in; and maybe the singer’s beloved, unlike his blood vessels, cannot really be in his arms. For what it is worth, I think Davidson (197x) was right to describe metaphorical usage as a kind of speech act, in which an expression is used (with the only meaning it has) to trigger cognitive processes beyond those that are plausibly involved in grasping any thought expressed with the sentence. But just as I wouldn’t say that querying is a matter of using a sentence that is true or false to ask a question, I wouldn’t say that “metaphoring” is a matter of using a sentence that is true or false to trigger cognitive processes that a hearer/reader might find interesting or aesthetically pleasing. Though singing (89a) is presumably a case of metaphorizing that leads a hearer to form a thought that can be recognized as obviously false—given sober reflection of a sort that makes it hard to enjoy a lyric, a horror movie (see Carroll 1990), and many other valuable things.

24 If ‘in’ has satisfiers, it can be used to specify them. Theorists might work out what the satisfiers would have to be, invent a language in which ‘In(_, _)’ has those satisfiers, and hypothesize that ‘in’ is truth-theoretically equivalent.
But given (C), it’s not easy to see how (53) could have a truth condition while (1) does not.

(C) each human language is an I-language in Chomsky’s sense

Other considerations regarding context sensitivity, illustrated with (69) and (87),

(69) The kettle is black.

(87) Kim said that James is thin.

make it even more plausible that sentences of a human language do not have truth conditions. Reflection on examples like (88) suggests that action reports do not have truth conditions,

(88) Al chased Theo.

even if the reports concern episodes of chasing as opposed to episodes of saying. Examples like (8) and (71) bolster the case that (53) and (69) are like (1) in not having truth conditions.

(8) Vulcan has an iron core.

(71) France is a hexagonal republic.

3.5 Contexts and Communication

It might seem that simple observations regarding human communication provide the basis for an argument in other direction. Stanley (200x) tries to defend his version of (D) in part by showing how it preserves the idea that “from an utterance of a sentence, one gains information about the world.” He offers a familiar kind of example that is instructive.

If Hannah utters to Esther the English sentence “There is some chocolate in the kitchen”, and if Esther wants chocolate, she will go to the kitchen. Even if Esther has never heard that particular sentence before, if she speaks English, she will understand what information it conveys about her physical environment (p. xx).

It’s not entirely clear, from this passage, just what Stanley takes the relevant explanandum to be: that sentences convey information; that one can gain information from sentential utterances; or that given a sentence, one can gain information from noting that it was uttered. The details may be important, given the differences between Davidsonian and Kaplanian strategies for accommodating context sensitivity, and corresponding debates about whether proposed truth theories for human languages should have theorems concerning utterances. It may also be important to be clear about the relevant notion of information, which is presumably not quite Shannon’s (1948) technical notion. I’m also not sure how Stanley’s point is supposed to generalize to the English sentences (69) and (89), in light of Travis’ (199x) discussion;

(89) There is some milk in the refrigerator.

see chapter zero. But I can grant that Hannah very likely conveyed some kind of information to Esther by pronouncing an instruction for how to build a thought.

If Esther speaks English and is suitably attentive, she will construct a thought in response to Hannah’s pronouncement; and Esther may reasonably believe that Hannah expected Esther to construct such a thought, that Hannah is not being cruel, etc. If ‘chocolate’ and ‘kitchen’ are lexical instructions for how to access atomic concepts, then hearing these words will lead speakers of English—i.e., those who have acquired an I-language that counts as an idiolect of English—to access similar concepts. And if Hannah knows what she is talking about, as opposed to just making stuff up, then Esther gets some information (as opposed to misinformation) about the location of some chocolate. But I don’t see any reason for thinking that the sentence itself is a source of this information as opposed to a tool for conveying it in complicated human way.

Again, the real issue is about how tightly meaning is connected to truth. Observations regarding human communication don’t settle this issue. Nor do they counterbalance the reasons for rejecting (D). Stanley goes on to say that it isn’t clear how one can construct a verifiable theory of meaning if one denies that sentences have truth conditions. I think this “only game in
town” argument is overrated. We can use speakers’ judgments—concerning implications, ambiguity, and whether or not contextualized uses of sentences are correct—as evidence for or against specific hypotheses about human I-languages, even if these expression-generating procedures are indirectly related to the truth conditions of truth-evaluable thoughts.

Stanley suggests that certain facts about communication can only be explained via theories according to which sentences bear a certain relation to information, in some sense that motivates truth-theoretic semantics. I disagree. If talk of information is standing in for talk of truth conditions, or a technical notion of proposition (cp. Soames 1987), then it is question-begging to assume that sentences bear any natural relation to information. If ‘information’ is being used in its ordinary way, then Hannah clearly conveyed information to Esther. But I don’t see why such facts can’t explained if meanings are instructions for how to build concepts with which humans can think about things in certain ways, often with the result that speakers and hearers think about the same aspects of a shared environment.

That said, Stanley’s appeal to unpronounced indices provides a possible answer to the question of how anything could be true given the context sensitivity of truth. This is important. For one might think that even if certain actions are paradigmatic examples of things that are true: (i) some sentences have to be true, at least relative to assignments of values to variables, on pain of nothing being true; (ii) truth has to be characterized indirectly, via satisfaction of sentences by assignments; but (iii) it’s not plausible that mental sentences exhibit the systematicity required to impose satisfaction conditions on sequences. So one might think that sentences of a human language are the best candidates we’ve got for sentences that have satisfaction conditions, at least if the context sensitivity of truth can be accommodated with many indices, in roughly the way that Stanley suggests. I’m inclined to agree with (i) and (ii); and I feel the pull of (iii). But my suggestion, developed in chapters five and six, will be that there are various kinds of mental sentences: those we share with other animals; those we can form by virtue of having a language faculty that lets us introduce lexical items, perhaps thereby introducing formally new concepts that combine systematically; and those we can form later, having already acquired a human language, when we engage in serious enquiry.

The existence of truths doesn’t show that sentences of a human I-language are among them. Many sentences in a math/logic/physics/biology/semantics textbook, or this chapter, are hybrids that mix vocabulary from a human language that a child could acquire with invented symbolism whose interpretation is stipulated in a way that only cognoscenti can follow. These hybrids are used in specialized contexts where it is understood that formulating truths—and hence, achieving the kind of precision required to formulate truths of the sort in question—is the goal, even if it is achieved only rarely and partially. Thoughts assembled via these hybrids are, in my view, better candidates for having Tarskian satisfaction conditions than the expressions generable by I-languages that children can acquire. But this is advertisement. Having digressed, in order to connect the details of this chapter with other considerations that bear on the question of how meaning is related to truth, let me return to some remaining details.

4. Back to Troublemakers
In this section, I consider a few attempts to reconcile thesis (D) with Liar Sentences.

(D) for each human language H, some theory of truth is the core of a correct theory of meaning for H

4.1 Hierarchy Revisited
Recall Lycan’s (2012) suggestion that a sentence can be a sentence of both a “core” language λ that contains no semantic vocabulary and an “expanded” language λ’ that contains such
vocabulary, including a truth predicate for λ. A more expanded language λ'' contains a truth predicate for λ'. And so on. On this view, (17) can be true-in-λ, and true-in-λ',

(17) Bessie is not blue.

but (1) is not a sentence of λ, and so not true-in-λ. Although (1) may be true-in-λ'.

(1) Lari is not true.

Initially, this seems like an attractive idea. But if λ and λ' are human I-languages, then it’s not clear how it is supposed to work. If λ and λ' both generate expression (17)—a particular articulation/meaning pair—and λ' generates (1) while λ does not, then it’s hard to see what λ can be apart from an arbitrarily restricted variant of λ': the same generative procedure, apart from somehow not having access to a few of the lexical items included in the specification of λ'. So it’s hard to see how a truth theory for the generative procedure λ' could, except by arbitrary stipulation, fail to be a truth theory for the generative procedure λ. We can, if we like, describe a lexicon L' and combinatorial system C as the result of adding lexical items to an I-language specified in terms of C and a smaller lexicon L. We can then introduce two extensionally distinct notions: true-in-CL and true-in-CL'. But it’s hard to see how this bears on the question of whether there is, independent of what semanticists choose to do, a Tarski-style theory of truth for the generative procedure characterized by the generative procedure CL'.

Advocates of (C) can agree that ‘English’ is not a name for any one human language.

(C) each human language is an I-language in Chomsky’s sense

Those who speak English have acquired I-languages that are very similar in many respects that matter for communication and various socio-political purposes. And a “monolingual” speaker of English is not limited to a single I-language. We can acquire new lexical items and use different “registers” in different conversations. In these senses, we can and do use distinct I-languages that massively overlap. But I see no reason, apart from fealty to (D) in the face of (C),

(D) for each human language H, there is a theory of truth

that is the core of a correct theory of meaning for H

for thinking that human I-languages exhibit natural “cuts” corresponding to Lycan’s hierarchy.

It can be tempting to slide into E-language thinking, characterize languages in terms of disposition to utter, and say that an utterance can be both true-in-English and true-in-German. But even if (59) and (60) sound alike, it doesn’t follow any I-Language generates both.

(59) Empedocles leaped.
(60) Empedocles liebt.

And if it turned out that the sound of ‘leaped/liebt’ had both meanings in two I-Languages, that would be the way to describe the discovery: in this respect, distinct procedures turned to be alike. But our theories of the languages acquired would remain theories of the procedures acquired—not theories that target (alleged) corresponding sets of articulation-meaning pairs and specify the meanings by assigning truth conditions to the relevant articulations relative to a syntax.

Lycan admits that his proposal is ad hoc in some respects. But his stated aim is to provide an alternative to Ludwig and Lepore’s (200x) suggestion that derivable biconditionals like (1-T)

(1-T) True(‘Lari is not true.’) ≡ ~True(Lari)

can be interpretive (i.e., correct meaning specifications) without being true, and that the axioms of an interpretive truth theory can yield contradictions given true assumptions. On this view, a truth theory for a human language H doesn’t have to be true to be the core of a correct theory of meaning for H. It’s good enough if the theory generates an interpretive T-theorem for each declarative sentence of H, but no uninterpretive theorems; see also Eklund (2002). I share Lycan’s suspicion of the idea that false truth theories can serve as correct meaning theories. This
seems like an option of penultimate resort, with embracing contradiction as the last resort. And perhaps Lycan’s defense of (D) is, all things considered, the better one. But I also sympathize with Ludwig and Lepore, who recognize that given Liar Sentences, (D) suggests that correct meaning theories will be based on false theories of truth.

In the end, my response to Ludwig and Lepore (and Lycan) will be that we can abandon (D) in favor of the claim that meanings are concept construction instructions. But here, let me note that while Ludwig and Lepore are not explicit about what they take human languages to be, they seem to assume that acquiring a spoken language is a matter of acquiring a capacity to interpret certain utterances. From this Davidsonian perspective, a truth theory for a human language \( H \) is part of a larger theory of action for a speaker of \( H \); where for these purposes, actions include episodes of speech and comprehension. If the theorist’s goal in formulating a truth theory is to ascribe it to speakers of \( H \)—as opposed to describing how \( H \) is related to a certain domain, about which the theorist wants to hypothesize by using \( H \)—then the formulated truth theory need not be true, much as beliefs ascribed to speakers need not be true. Indeed, it can be appropriate to ascribe an incorrect truth theory for a language to a person who has made a certain kind of mistake. But this suggests that the person is trying to use a language whose character is determined independently of that person. And whatever such a language is, it is presumably not an I-language that the person already implements.\(^{25}\)

If a child thinks that ‘+’ signifies addition, and infers that ‘8 + 2 = 10’ is true, then it may be appropriate to explain the child’s endorsement of this sentence by ascribing to the child a partly false theory of how the arithmetic notation is to be understood. Dummett (197x, 198x) suggested that any speaker could likewise be wrong about certain aspects of her naturally acquired human language. On his view, I might be wrong about what ‘livid’ means in my own language—and not just wrong about how certain “experts” would use a homophonous word, or wrong about certain conventions governing public discourse. Davidson (1986) rejected this suggestion, along with the underlying idea that a speaker’s native language is independent of the speaker’s grasp of that language. As an illustration, Davidson said that when Mrs. Malaprop utters ‘That is a nice derangement of epitaphs’, her sentence (and not just she) means what our sentence ‘That is a nice arrangement of epithets’ means. On this view, homophonic translation of Mrs. Malaprop would be incorrect. Her utterance is true if and only if the demonstrated thing is a nice arrangement of epithets; and if she pragmatically implies anything else, that is in part because her utterance has this (non-homophonically reported) truth condition.

That debate about the possibility of “linguistic error” (e.g., that ‘epitaph’ is true of epithets) focused on specific lexical meanings. It is far more radical to abstract from such errors and ascribe a false truth theory to all speakers of English—especially in the absence of a true truth theory and a plausible account of the facts (psychological, social, causal, or whatever) in virtue of which such a theory would be true. Given that human languages include “semantic vocabulary,” which speakers use as they do, we may find ourselves committed to contradictory policies concerning how expressions of a human language can be used to make true claims. That is one way of describing the Liar Paradox. But if only because we don’t know which of our commitments are false, the contradictory character of these commitments is no justification for doubling down on (D), and concluding that false truth theories can be correct meaning theories.

\(^{25}\) Though perhaps it is a “\( \Psi \)-language” in Ludlow’s (2011) externalistic sense. One can stipulate that I-languages are “private languages”—expressions of which do not have “objective” correctness conditions—and hence not “real languages.” But then human languages may not be real in the stipulated sense.
(D) for each human language $H$, there is a theory of truth that is the core of a correct theory of meaning for $H$

This leaves us with no clear idea of what the hypothesized false truth theories are theories of.

One can take an instrumentalist attitude towards truth theories, treating them as mere calculation devices that pair object language expressions with “translations” in a certain formalism. But even if one can show that the formal translations are independently attractive as meaning specifications—and this is no small task—that would leave the question of whether the appeal to truth plays any important role in the explanation of why human linguistic expressions mean what they do. Why not instead view the object language expressions as instructions for how to assemble thoughts of the sort depicted with the metalanguage formalism?

### 4.2 False Theories and I-languages

One might, however, try to combine the spirit of Ludwig and Lepore’s suggestion—and at least some of Davidson’s reply to Dummett—with thesis (C) by adopting (BF).

(C) each human language is an I-language in Chomsky’s sense

(BF) each human language $H$ is an I-language such that some false

truth theory for $H$ is the core of a correct theory of meaning for $H$

One might say that Mrs. Malaprop not only has an I-language with a lexical item ‘epitaph’ that is true of epithets, she has and endorses a truth theory (for this I-language) that includes a true axiom whose content is that ‘epitaph’ is true of epithets. On this view, Mrs. Malaprop is right about her word, not merely wrong about ours. But the idea would be that each speaker of a human language $H$ can be wrong about the truth conditions of certain sentences of $H$, at least if we each have and endorse a truth theory whose theorems include biconditionals like (1-T).

(1-T) $\text{True}(\text{’Lari is not true.’}) \equiv \neg \text{True(Lari)}$

This suggestion can be tempting if one views children as interpreters, trying to figure out the truth conditions of alien utterances. One imagines a little field linguist with a concept of truth that gets “overextended” by a simple theory, with the result that the child is apt to draw incorrect conclusions like (1-T), as a side effect of being able to draw correct conclusions like (17-T).

(17-T) $\text{True}(\text{’Bessie is not blue.’}) \equiv \neg \text{Blue(Bessie)}$

But this is to view the quoted sentence as an expression of the language used by the speakers who are being interpreted, not as an expression generated by the interpreter’s own I-language. And our goal is not merely to explain why each us comes to think—perhaps wrongly—that the sentences used by others have certain truth conditions. Given a certain I-language (or a family of I-languages that are the same in certain respects), our goal is to explain why the expressions of that I-language (family) mean what they do. But if we want to know what the semantic properties of (17) are, and how they are compositionally determined,

(17) Bessie is not blue.

it doesn’t help to be told that (17-T) follows from some interpreter’s false theory.

One can say that (17-T) follows from a false but interpretive theory of truth for English. But this reminds us that the truth theory allegedly represented by ordinary speakers is governed by a very weak logic, making the theory impervious to logical pressure.26 Once again, this raises

---

26 When theorizing about a cognitive competence whose application to cases is subject to logically contingent constraints, we need to distinguish (at least) two notions of implication. One concerns what a theory implies given the rest of science, logic and mathematics very much included. Ideally, one wants to make all the implications of a theory manifest, whatever implications turn out to be; see, e.g., Frege (1879, 188x, 189x., 189x). So we want a logic that is more powerful than the propositional calculus. And for purposes of figuring out what follows from a theory, mathematical implications usually count, even though not even arithmetic—much less analysis or geometry—is

33
the question of whether the appeal to truth plays any important role in the explanation of why human linguistic expressions mean what they do, and why this “error theory” is preferable to viewing human linguistic expressions as concept assembly instructions. Advocates of (BF) need to say which false implications of a hypothesized truth theory Θ do not tell against the hypothesis that Θ is the core of meaning theory for a given human language. And again, Liar Sentences are not the only barriers to providing true theories of truth for human languages.

That said, let me conclude this section by noting my appreciation for Lycan, Ludwig and Lepore’s recognition that Liar Sentences present a serious challenge for advocates of (D).

(D) for each human language H, there is a theory of truth that is the core of a correct theory of meaning for H

In my view, their proposals honestly reflect the difficulties and the substantive ancillary assumptions needed to accommodate sentences like (1).

(1) Linus is not true.

In retrospect, Davidson’s (1967a) discussion seems overly brief and optimistic.

The semantic paradoxes arise when the range of the quantifiers in the object language is too generous in certain ways. But it is not really clear how unfair to Urdu or to Wendish it would be to view the range of their quantifiers as insufficient to yield and explicit definition of ‘true-in-Urdo’ or ‘true-in-Wendish’....In any case, most of the problems of general philosophical interest arise within a fragment of the natural language that may be conceived as containing very little set theory (28-29).

Restricting the range (or domain) of human language quantifiers doesn’t help with (1). I don’t know how Davidson was counting “problems of general philosophical interest,” But one interesting question is whether (D) is true. And in thinking about other reasons for being suspicious of truth-theoretic semantics, we shouldn’t assume that examples like (1) present mere technical problems for a proposal that shines with regard to other examples.

---

reducible to “pure logic.” By contrast, Foster’s Problem suggests a far more parochial notion of implication, corresponding to the constrained capacities of ordinary speakers to deploy semantic competence in ways that support the comprehension of boundlessly many novel expressions. As with other natural competences, deployment is constrained by the relevant representational forms (which are biologically implemented), and so unlikely to be closed under whatever forms of deduction a theorist uses to evaluate theories. Indeed, the constraints on an interpretive truth theory are so severe as to raise the question of whether the theory really specifies truth conditions for sentences, especially if it is granted that the conjunction of the axioms is false.
Chapter Five: Massively Monadic, Potentially Plural

In this chapter, I return to the idea that lexicalization is a process in which available concepts are used to introduce concepts that can be conjoined in certain ways.

1. Introduction and Foreshadowing

In chapter one, I posited two combinatorial operations, M-junction and Θ-junction. Recall that only monadic concepts can be M-joined: combining φ(_) with ψ(_) in this way yields another monadic concept, \( φ(_) ∨ ψ(_) \), which applies to whatever both \( φ(_) \) and \( ψ(_) \) apply to. But a dyadic concept can be Θ-joined with a monadic concept: combining \( θ(_-,-) \) with \( φ(_) \) in this way yields the monadic concept \( θ(_-, -) ∨ φ(_) \); where the internal/right slot of \( θ(_-, -) \) is linked to the unsaturated position of the adjacent monadic concept, which is also the position that \( ∃ \) closes. So \( θ(_-, -) ∨ φ(_) \) is a concept of things that bear the dyadic relation—perhaps a thematic relation that events bear to certain participants—to at least one thing that falls under the concept \( φ(_) \). In my view, a phrase formed by combining two lexical items is an instruction to fetch a corresponding pair of concepts and join them in one of these two ways. If this is correct, then concepts fetched via lexical items must be monadic or dyadic, even if this requires that fetchable concepts be introduced as opposed to being merely labeled in the process of lexicalization. I think there is considerable evidence for this view. More generally, I think lexicalization is like Fregean definition in being a cognitively creative process, though without the hierarchy of conceptual types; see §§1.4 and 1.5 of chapter two.

1.1 Junctions without Variables

The operation of M-junction permits construction of concepts like \( \text{red}(\_)^\text{coat}(\_). \) As discussed in chapter one, such concepts seem to be governed by a natural logic that typically licenses conjunct-reduction, but licenses conjunct-expansion in certain special contexts like the scope of negation. As the medieval logicians knew—and as Frege’s work highlighted—a mere logic of predicates is inadequate even for natural thought, which involves at least some relational notions. But given just a few atomic dyadic concepts, Θ-junction makes it possible to build boundlessly many concepts that have dyadic constituents, as in \( \text{stab}(\_)^\text{red}(\_)^\text{coat}(\_) \) and \( \text{stab}(\_)^\text{red}(\_) \). The conjunction \( \text{stab}(\_)^\text{red}(\_)^\text{coat}(\_) \) can be seen as a concept of a man’s red coat, and \( \text{stab}(\_)^\text{red}(\_) \) as a concept of a man’s red coat. In my view, a phrase formed by combining two linguistic items is an instruction to fetch a corresponding pair of concepts and join them in one of these two ways. If this is correct, then concepts fetched via lexical items must be monadic or dyadic, even if this requires that fetchable concepts be introduced as opposed to being merely labeled in the process of lexicalization. I think there is considerable evidence for this view. More generally, I think lexicalization is like Fregean definition in being a cognitively creative process, though without the hierarchy of conceptual types; see §§1.4 and 1.5 of chapter two.

1.1 Junctions without Variables

The operation of M-junction permits construction of concepts like \( \text{red}(\_)^\text{coat}(\_). \) As discussed in chapter one, such concepts seem to be governed by a natural logic that typically licenses conjunct-reduction, but licenses conjunct-expansion in certain special contexts like the scope of negation. As the medieval logicians knew—and as Frege’s work highlighted—a mere logic of predicates is inadequate even for natural thought, which involves at least some relational notions. But given just a few atomic dyadic concepts, Θ-junction makes it possible to build boundlessly many concepts that have dyadic constituents, as in \( \text{stab}(\_)^\text{red}(\_)^\text{coat}(\_) \) and \( \text{stab}(\_)^\text{red}(\_)^\text{coat}(\_). \) The conjunction \( \text{stab}(\_)^\text{red}(\_)^\text{coat}(\_) \) can be seen as a concept of a man’s red coat, and \( \text{stab}(\_)^\text{red}(\_) \) as a concept of a man’s red coat. In my view, a phrase formed by combining two linguistic items is an instruction to fetch a corresponding pair of concepts and join them in one of these two ways. If this is correct, then concepts fetched via lexical items must be monadic or dyadic, even if this requires that fetchable concepts be introduced as opposed to being merely labeled in the process of lexicalization. I think there is considerable evidence for this view. More generally, I think lexicalization is like Fregean definition in being a cognitively creative process, though without the hierarchy of conceptual types; see §§1.4 and 1.5 of chapter two.

1.1 Junctions without Variables

The operation of M-junction permits construction of concepts like \( \text{red}(\_)^\text{coat}(\_). \) As discussed in chapter one, such concepts seem to be governed by a natural logic that typically licenses conjunct-reduction, but licenses conjunct-expansion in certain special contexts like the scope of negation. As the medieval logicians knew—and as Frege’s work highlighted—a mere logic of predicates is inadequate even for natural thought, which involves at least some relational notions. But given just a few atomic dyadic concepts, Θ-junction makes it possible to build boundlessly many concepts that have dyadic constituents, as in \( \text{stab}(\_)^\text{red}(\_)^\text{coat}(\_) \) and \( \text{stab}(\_)^\text{red}(\_)^\text{coat}(\_). \) The conjunction \( \text{stab}(\_)^\text{red}(\_)^\text{coat}(\_) \) can be seen as a concept of a man’s red coat, and \( \text{stab}(\_)^\text{red}(\_) \) as a concept of a man’s red coat. In my view, a phrase formed by combining two linguistic items is an instruction to fetch a corresponding pair of concepts and join them in one of these two ways. If this is correct, then concepts fetched via lexical items must be monadic or dyadic, even if this requires that fetchable concepts be introduced as opposed to being merely labeled in the process of lexicalization. I think there is considerable evidence for this view. More generally, I think lexicalization is like Fregean definition in being a cognitively creative process, though without the hierarchy of conceptual types; see §§1.4 and 1.5 of chapter two.
concepts that can be M-joined or Θ-joined, even if such concepts have to be introduced. So evidence that lexicalization is such a process can support the proposed conception of composition, at least if many prelinguistic concepts are like \textsc{gave}(\_,\_,\_) and \textsc{bessie} in not being joinable via the restrictive operations. In this chapter, I provide such evidence. Chapter six is devoted to compositional details and the more general claim that all phrasal meanings are instructions for how to assemble conjunctive monadic concepts.

On this view, lexical items can be used to fetch dyadic concepts, but composition preserves only one unsaturated position. In terms of the Fregean hierarchy of concepts: phrases correspond to concepts of types \textlangle e, t\textrangle; and the lexicon of a human language is only slightly less restricted, permitting access to some concepts of type \textlangle e, et\textrangle. But if meanings are not used to access or build concepts of types \textlangle e\rangle or \textlangle e, t\textrangle, with talk of entities and truth values being otiose, we may as well call the two basic semantic types \textlangle M\rangle and \textlangle D\rangle for ‘monadic’ and ‘dyadic’. And if all phrases correspond to concepts of type \textlangle M\rangle, then not only is there no hierarchy of non-basic types of the form \textlangle α, β\textrangle, there are no non-basic types. That said, my proposal allows for (and requires) lexical items and phrases corresponding to concepts that are plural—and quantification that is second-order—in Boolos’ (1998) sense. Indeed, the idea developed in this chapter and the next is that human I-languages let us introduce and systematically combine concepts that can be viewed as composable symbols of a second-order extension of a massively monadic fragment of the (first-order polyadic) language, PL, that was described in chapter two.

1.2 Introducing vs. Labeling

In chapter one, I readily granted that humans have relational concepts like \textsc{chase}(\_,\_) and \textsc{give}(\_,\_,\_); where the dyadic concept applies to ordered pairs \textlangle α, β\textrangle such that (at some time t) α chases β, and the triadic concept applies to triples \textlangle α, β, γ\textrangle such that α gives β to γ. But it doesn’t follow that \textsc{chase}(\_,\_) is the meaning of ‘chase’, or that this word meaning is dyadic.

For reasons discussed in chapter three, I assume that verbs are semantically associated with “event positions” of some kind, even if their relation to mind-independent events remains unclear. But this assumption is compatible with many specific claims about how words of a human I-language combine, and how verb meanings are related to concepts and to the meanings of sentences like (1), in which ‘chase’ combines with more than one grammatical argument.

\begin{itemize}
  \item (1) A girl saw a dog chase a cow.
\end{itemize}

One can hypothesize that ‘chase’ is true of ordered triples \textlangle e, α, β\textrangle such that e is a chase by α of β—or in overtly mentalistic terms, that ‘chase’ accesses a triadic concept with slot for events. And perhaps ‘give’ accesses a tetradic concept that applies to \textlangle e, α, β, γ\textrangle iff e is a giving by α of β to γ. This assumes an operation of saturation, or its equivalent, as the correlate of combining a verb (or verb phrase) with a grammatical argument. But another possibility is that both ‘chase’ and ‘give’ access monadic concepts of events, with grammatical arguments corresponding to concepts like \textsc{∃}[\textsc{agent}(\_,\_)^\textsc{dog}(\_)] and \textsc{∃}[\textsc{patient}(\_,\_)^\textsc{cow}(\_)]. Given Θ-junction, it is also possible that verbs access dyadic concepts like \textsc{passive-chase}(\_,\_) and \textsc{passive-give}(\_,\_); where such concepts apply to event-individual pairs (see Kratzer 199x). This permits concepts like \textsc{∃}[\textsc{agent}(\_,\_)^\textsc{dog}(\_)]^\textsc{∃}[\textsc{passive-chase}(\_,\_)^\textsc{cow}(\_)]. But if combining a verb with an argument always calls for Θ-junction, this constrains hypotheses about the meaning of ‘chase’; \textsc{chase-by-of}(\_,\_,\_) cannot be Θ-joined with any concept.

In this sense, it isn’t obvious which Fregean type ‘chase’ corresponds to, even in a specific construction like (1): \textsc{chase}(\_) is of type \textlangle e, t\textrangle; \textsc{passive-chase}(\_,\_) is of type \textlangle e, et\textrangle; \textsc{chase-by-of}(\_,\_,\_) is of type \textlangle e, e, et\textrangle. But it seems unlikely that in (1), ‘chase’ is used to access a concept of chaser-chasee pairs. Perhaps children initially connect the pronunciation of
‘chase’ with \text{CHASE}(\_, \_). But if so, then connecting the pronunciation with the meaning of ‘chase’ requires further work. Similarly, if combining ‘chase’ with a grammatical argument calls for \Theta-junction, then acquiring the lexical item requires more than connecting the pronunciation with \text{CHASE-BY-OF}(\_, \_, \_). If children simply acquire and label \text{PASSIVE-CHASE}(\_, \_) or \text{CHASE(\_)}, they don’t need to \textit{introduce} some such concept that has no slot for chasers. That is compatible with my proposal. But then we need to explain the intuition that every chase has an agent.

If combining a verb with an argument always calls for \Theta-junction, this also constrains hypotheses about the meaning of a name like ‘Bessie’; \text{BESSIE} cannot be \Theta-joined with any concept. But in any case, the available combinatorial operations constrain the options regarding which concepts can be accessed with words. And if only for this reason, it seems unlikely that lexicalization is always a process in which an available concept is simply labeled with an item that inherits its type from the labeled concept. It seems that any plausible view of composition will require a dovetailing conception of lexicalization according to which some concepts are used to introduce formally distinct but analytically related concepts.

One might have hoped for the following transparent mapping from conceptual types to linguistic types: lexicalizing a denoting concept yields a proper name of type \langle e \rangle; lexicalizing a monadic concept yields a noun or intransitive verb (or adjectival/adverbial expressions) of type \langle e, t \rangle; lexicalizing a dyadic concept yields a transitive verb, of type \langle e, et \rangle; etc. Ignoring tense for simplicity, one might have expected ‘chase’ and ‘Bessie’ to be labels for \text{CHASE}(\_, \_) and \text{BESSIE}, with ‘chase Bessie’ indicating \text{CHASE}(\_, \text{BESSIE}); where this partly saturated concept applies to an individual \alpha iff (at some time \text{t}) \alpha \text{chases Bessie}. Analogizing to 19th century chemistry, one could hypothesize that ‘chase’ inherits a valence of −2 from the concept lexicalized, ‘Bessie’ inherits a valence of +1, and so ‘chase Bessie’ has a valence of −1; cp. the compound Hydroxide, \langle \text{O}^{2-}\text{H}^{+} \rangle. This would capture the idea that ‘chase Bessie’ is like the intransitive verb ‘bark’, which can combine with a single argument to form a complete clause, much a Hydroxide ion can combine with a Hydrogen atom to form a stable molecule of water.

As we’ll see, there is more than one way of understanding the valence analogy. But for now, think of a verb’s valence as a (hypothized) \textit{structural} property of that verb—akin to the adicity of an unsaturated concept. The familiar picture is that a verb—or more generally, a lexical predicate—provides a “sentence frame” in much the way that an unsaturated concept provides a “thought frame”; see chapter one. Correlatively, one might say that names—or more generally, referential devices—can saturate lexical predicates, each of which “takes” a certain number of \textit{grammatical} arguments, perhaps subject to certain typological or “selectional” restrictions. On this view, Frege was wrong to think that human linguistic expressions exhibit subject-predicate structure \textit{as opposed to} predicate-argument structure; expressions are said to reflect both the saturated/saturating contrast and adicity contrasts among unsaturated concepts.

As noted in chapters one and two, this simple idea does not yet accommodate quantificational constructions like ‘chased every cow’, given the limits of saturation (as opposed to Church-style function-application) as a composition operation. But for these purposes, one might want to bracket questions about quantification and ignore the formal differences between \text{CHASE}(\_, \_) and \lambda x. \lambda x'. \text{T} \equiv \text{CHASE}(x', x). So let’s concessively extend the notion of labeling to include the additional apparatus required to construct \lambda x. \lambda x'. \text{T} \equiv \text{CHASE}(x', x) from \text{CHASE}(\_, \_). This relatively minimal re-presentation of a conceptual content preserves the concept’s original “argument structure;” likewise for recoding \text{GIVE}(\_, \_) as \lambda x. \lambda x'. \lambda x''. \text{T} \equiv \text{GIVE}(x'', x', x). But it isn’t mere labeling if a mind initially connects the pronunciation of ‘chase’ with \text{CHASE}(\_, \_) but then uses this dyadic concept to introduce \lambda x. \lambda x'. \lambda x''. \text{T} \equiv \text{CHASE-BY-OF}(x'', x', x) and
connect this triadic concept with the pronunciation of ‘chase’.

Creating a concept of higher adicity, in ways that allow for conjunctive adverbial modification, reflects a more significant departure from the initial concept. It also reflects a capacity to go up the Frege-hierarchy, using a concept of type \(<e, et>\) to introduce one of type \(<e, e, et>\). If a mind initially connects the pronunciation of ‘chase’ with \(\text{CHASE}(\_ , \_ )\) but then uses this concept to introduce concept to introduce \(\lambda x . \lambda x'. T = \text{PASSIVE-CHASE} (x, x')\) or \(\lambda x . T = \text{CHASE}(x)\), that is also interesting. Of course, evaluating any such claim about lexicalization is hard, if only because it is unclear which concepts are natural for children independent of acquiring words. To understate, we don’t have good theories of which concepts are available to children as natural candidates for initially labeling with pronunciations. But some familiar generalizations tell against the idea that lexicalization is merely a process of labeling concepts, and in favor of the idea that at least often, lexicalization is a process in which available concepts are used to introduce formally new concepts that can be M-joined or \(\Theta\)-joined.

These same considerations may also tell against appeals to “grammatical argument structure,” depending on how that vexed notion is understood; see Borer (2005), cp. Ramchand (2008); Lohndal (2014). But in the next section, I just state the generalizations somewhat starkly. Section three offers clarifications and qualifications, along with discussion of some implications. In section four, I extend the notion of concept introduction beyond cases of new adicities, and suggest that lexicalization is a process in which available concepts are used to introduce concepts that are “number neutral”—neither singular nor plural, neither count nor mass. This will connect back to reasons for analyzing verbs in terms of monadic predicates of events, following Schein (1993, 2002, forthcoming), who drew on Boolos’ (1998) account of second-order quantification as plural quantification. This will also prepare the way for chapter six. But the next order of business is to argue against the idea that lexicalization is merely a process of labeling concepts.

2. Four Generalizations

In this section, I note four generalizations, each of which is familiar but often set aside as a complication as opposed to a telling symptom of lexicalization and composition in human languages. First, while humans have denoting concepts, proper nouns seem not to be lexical items of type \(<e>\), textbook treatments to the contrary notwithstanding. On the contrary, the lexical noun ‘Bessie’ seems to fetch a monadic concept, perhaps one that applies to things called ‘Bessie’; though the atomic noun can combine with a functional item to form a name that corresponds to a complex concept that applies (in contexts) to only one individual. Second, unsaturated concepts seem to be lexicalized with words that are often flexible about how many arguments they combine with in sentences. And the pattern of flexibility suggests that these words are used to fetch monadic or dyadic concepts. Third, “supradyadic” concepts like \(\text{GIVE}(\_ , \_ , \_ )\) are not lexicalized with verbs that take more than two grammatical arguments.

Conceptually essential participants—e.g., the recipients in events of giving (or donating)—are often represented with prepositional phrases that are grammatically optional. Fourth, there is a constraint on how the subject/object asymmetry is related to the agent/patient distinction, as if concepts that have an unsaturated position corresponding to agents are lexicalized with words that subsequently fetch related but distinct concepts that do not have such a slot.

2.1 Limited Relations

This last generalization is often discussed in other contexts. Transitive constructions respect a substantive constraint on how “thematic order” can be expressed.² Presumably, ‘chase’

² See Gruber (1965), Fillmore (1968), Dowty (1991), Baker (1997), and further references there.
lexicalizes a relational concept. But even if this concept already has a slot for events, that doesn’t explain why in (1), ‘a dog’ corresponds to the chaser of the chasing.

(1) A girl saw a dog chase a cow.

We can assume that ‘a dog chase a cow’ has the following constituency structure, with ‘a cow’ as the internal argument of ‘chase’: [[a dog] [chase [a cow]]]. But this still doesn’t explain why ‘a cow’ corresponds to the thing chased. Likewise, (2) implies that was Juliet was a kisser,

(2) Everyone saw Juliet kiss Romeo.

but not that Romeo was.

This thematic contrast might be described by saying that ‘chase’ corresponds to the concept indicated with (1a) as opposed to (1b), and likewise for ‘kiss’;

(1a) $\lambda x'' . \lambda x' . \lambda x . T \equiv \text{CHASE-BY-OF}(x, x'', x')$

where by stipulation, \text{CHASE-BY-OF}(_, _, _) applies to $<\epsilon$, $\alpha$, $\beta>$ if $\epsilon$ is a chase by $\alpha$ of $\beta$, and order of function-application tracks order of grammatical arguments. On this view ‘Fido chase Bessie’ corresponds to $\lambda x . T \equiv \text{CHASE-BY-OF}(x, \text{Fido}, \text{Bessie})$; where this function maps $\epsilon$ to $T$ iff $\epsilon$ is a chase by Fido of Bessie. But this just encodes the contrast. We could have stipulated that \text{CHASE-OF-BY}(_, _, _) applies to $<\epsilon$, $\alpha$, $\beta>$ if $\epsilon$ is a chase of $\alpha$ by $\beta$. In which case, (1b)

(1b) $\lambda x'' . \lambda x' . \lambda x . T \equiv \text{CHASE-OF-BY}(x, x', x'')$

is equivalent to (1b), not (1a). And we have no reason for thinking that \text{CHASE-OF-BY}(_, _, _) is less natural, as a concept, than \text{CHASE-BY-OF}(_, _, _).

Put another way, we can imagine a verb ‘asechay’ that corresponds to (1b). So we can imagine a language such that ‘Bessie’ is the direct object of the verb in (3),

(3) Fido asechayed Bessie.

but this imagined sentence means that there was a chase by Bessie of Fido. Similarly, we can imagine a language Glisheng, in which the pronunciation of ‘chase’ is associated with the concept \text{CHASE-OF-BY}(_, _, _). In the Glisheng sentence (4),

(4) Fido chased Bessie.

‘Bessie’ is the direct object of the verb and so corresponds to the chaser-role, while ‘Fido’ corresponds to the chasee-role. So we face the question of why human children regularly acquire verbs that are thematically like ‘chase’ as opposed to ‘asechay’. Or put another way, we face the question of why human children regularly acquire languages like English as opposed to Glisheng.

One can describe the generalization by elaborating (1a) and (1b) as (1aa) and (1bb).

(1aa) $\lambda x'' . \lambda x' . \lambda x . T \equiv \text{AGENT}(x, x') \& \text{CHASE}(x) \& \text{PATIENT}(x, x'')$

(1bb) $\lambda x' . \lambda x'' . \lambda x . T \equiv \text{AGENT}(x, x') \& \text{CHASE}(x) \& \text{PATIENT}(x, x'')$

Then one can say that the external argument of ‘kiss’ is associated with \text{AGENT}(_, _) while the internal argument is associated with \text{PATIENT}(_, _); cp. Parsons (1990) on “subatomic” semantics. Alternatively, one can replace \text{CHASE}(x) \& \text{PATIENT}(x, x'') with \text{PASSIVE-CHASE}(x, x''). But either way, this still leaves the question of why the pronunciation of ‘kiss’ is connected with the first triadic meaning as opposed to the second. One possible answer is that while verb meanings can be triadic, meanings like (2bb) are prohibited. Perhaps human composition operations can generate the meaning of (3), but while (2aa) is a possible word meaning, (2bb) is not. I don’t think there is a plausible source for any such thematic constraint on lexical items. So my suspicion is that neither (2aa) nor (2bb) can be a human word meaning: ‘chase’ fetches the monadic concept \text{CHASE}(_), or perhaps the dyadic concept \text{PASSIVE-CHASE}(_, _); ‘asechay’ would fetch the same concept, not any triadic concept with a slot for chasers. And if that’s right, I think it tells against the idea that verbs simply label the concepts they lexicalize.
On my view, verbs do not have thematically elaborable meanings. So the order of the chaser/chasee slots in $\text{CHASE}(\_ , \_)$ or $\text{CHASE}(\_ , \_)$ is irrelevant. Indeed, we need not assume that all polyadic concepts impose an order on their arguments. At least some relational concepts may be more like hubs with spokes; a conceptual hub may associate each of its spokes with a particular role, though not in any canonical order. In any case, whatever the nature of the concept lexicalized with ‘chase’, a child may associate this lexical item with a (perhaps introduced) concept that has a slot for chases, no slot for chasers, and perhaps no slot for chasees. If ‘chase’ corresponds to $\text{CHASE}(\_)$, the verb phrase [chase [a cow]] corresponds to the monadic concept of events $\exists[\text{PATIENT}(\_ , \_), \text{COW}(\_)]$. If ‘chase’ corresponds to $\text{PASSIVE-CHASE}(\_)$, then [chase [a cow]] corresponds to $\exists[\text{PASSIVE-CHASE}(\_), \text{COW}(\_)]$. Either way, the phrase corresponds to a concept of cow-chasings; and while such events have agents, neither $\Theta$-junction reflects this fact. But one can say that the verb phrase [a dog [chase [a cow]]] corresponds to $\exists[\text{AGENT}(\_ , \_), \text{DOG}(\_), \text{CHASING-OF-A-COW}(\_) ];$ where the second conjunct is spelled out in terms of $\text{CHASE}(\_)$ or $\text{PASSIVE-CHASE}(\_).$ 3

The familiar idea is that one way or another, the external argument of ‘chase’ is associated with the concept $\text{AGENT}(\_ , \_)$. That still leaves the question of why children don’t acquire languages that associate the external argument of ‘chase’ with $\text{PATIENT}(\_ , \_)$. Saying that ‘chase’ corresponds to $\text{PASSIVE-CHASE}(\_)$ begs this question; let $\text{UNERGATIVE-CHASE}(\_)$ apply to $<e, \alpha>$ iff $e$ is a chase by $\alpha$. I suspect that it is not arbitrary that internal arguments are not associated with $\text{AGENT}(\_ , \_)$. But even if it is, we can at least capture the relevant generalizations by saying that certain grammatical relations are associated with certain thematic relations. By contrast, it’s hard to see how to even report the generalizations if verbs can have meanings that are triadic, as opposed to at most dyadic.

As briefly noted in chapter two, there seems to be an analogous constraint on sentences formed by combining a quantificational determiner like ‘every’ with two predicates as in (5).

(5) Every cow is brown.

Presumably, ‘every’ lexicalizes some relational concept—perhaps $\text{INCLUDE}[\Psi(\_), \Phi(\_) ]$, with $\Phi(\_)$ corresponding to ‘cow’—that yields a complete thought when saturated by two monadic concepts. One can imagine a word ‘ryev’ that is otherwise like ‘every’, except for expressing the concept $\text{SUBSET}[\Psi(\_), \Phi(\_) ]$. Then (6) would have the meaning of (7).

(6) Ryev cow is brown.

(7) Every brown thing is a cow.

But as reviewed in chapter six, there is no such determiner. I think this is because pace standard claims, determiners do not fetch concepts of second-order relations; see Pieterski (2005a, 2006). But in any case, I will be highlighting the absence of words that might be expected if lexicalization were simply a matter of labeling available concepts.

To take another example, why do we construct phrases like ‘is bigger than Fido’? Why don’t we label $\text{BIGGER}(\_ , \_)$ with a monomorphemic transitive verb, and use sentences like (8)?

(8) *Sadie bigs Fido.

Or put another way, why don’t we use more morphemes to say what we say with (9)?

---

3 On the former view, verbs are treated like common nouns in terms of valence. On the latter, verbs are treated more like dyadic prepositions. For now, I don’t want to take a hard line on this issue; see Svenonius (2010) for relevant crosslinguistic discussion. I return to the issue in chapter six. Appeal to $\text{PASSIVE-CHASE}(\_ , \_)$ is compatible with a saturationist conception of composition according to which ‘chase Bessie’ corresponds to $\text{PASSIVE-CHASE}(\_ , \_ BESSIE)$. But the idea that ‘Bessie’ is of type $<e>$ may not be sustainable, given facts about names; see §2.4 below. So I don’t think that this is a reason for favoring $\text{PASSIVE-CHASE}(\_ , \_)$.
(9) Sadie is big.
One can imagine a language in which (9) and (10) are ungrammatical, but (8) and (11) are fine.
(10) Sadie is a big mare.
(11) *Sadie bigs most mares.

So assuming that the concept lexicalized with ‘big’ is somehow relational, one wants to know
why (9) and (10) are even possible, if relational concepts are typically lexicalized with predicates
that take two or more arguments in sentences. One can posit covert morphemes. But this is to
grant that the relational concepts are not simply labeled with open class items.

More generally, it seems that many “relational thoughts” are expressed with words that
fetch monadic concepts. If only because (12) implies (13),
(12) Trees dropped leaves onto the ground, as a student dropped her books.
(13) Leaves dropped onto the ground.

there is pressure to say that the lexical item ‘drop’ fetches a concept that has no unsaturated slot
for droppers, and that the transitive verb in (12) is formed by adding a covert morpheme to the
intransitive verb in (13). But it seems less plausible that ‘drop’ lexicalized a concept with no slot
for droppers, and that our initial concept of one thing dropping another has a monadic component.
Thinking about dropping seems to differ from thinking about falling, in that the former seems to
involve thinking about actions that have agents.

Relatedly, the concept of one thing felling another may have constituents that include a
concept of falling. Perhaps in this case, conceptual structure reflects morphological structure,
with the verb in (14) having the verb in (15) as a constituent; cp. Fodor and Lepore (200x).
(14) Lumberjacks felled trees in the forest.
(15) Trees fell in the forest.

But if DROP(_, _) is lexicalized with a “root” expression that does not itself take two arguments,
so that using the lexical item to express a relational thought requires an extra morpheme, one
wants to know why. My proposal is that ‘drop’ lexicalizes a relational concept that gets used to
introduce DROP(_, which can be M-joined with concepts like Ǝ[AGENT(,,)^STUDENT(,)].

2.2 Never Three or More
Constraints on the expression of supradyadic concepts seem to be even more severe.

While (16) and (17) are possible, (16a) and (17a) are also good sentences.
(16) I bet you ten dollars that Fido is a mutt.
(16a) I bet ten dollars.
(17) I sold her a horse for ten dollars.
(17a) I sold a horse.

So prima facie, neither ‘bet’ nor ‘sold’ requires four arguments: ‘that Fido is a mutt’ may specify
the content of a bet without being a grammatical argument of ‘bet’, even if the propositional
content of a bet is conceptually essential to it, just as ‘for ten dollars’ may specify the amount for
which a horse was sold without being a grammatical argument.4 One can posit covert arguments
in (16a) and (17a). But ‘her’ seems to be an optional constituent in (17). And with present tense,
(16) implies an offer—not yet uptake—raising further questions about the role of ‘you’.

I can’t prove that no verbs take four arguments, two of which are often covert. But if such
verbs are possible, in human I-languages, one wants to know why they are not more common.

4 See Munro (1982) on the adjunctive character of complementizer phrases. There is also evidence that verbs like
speech act (and propositional attitude) verbs fetch monadic concepts of event/states that have contents, as dyadic
concepts of individual-content pairs; see Pietroski (2000c), Motomura (2003), Moltmann (201x).
For plausibly, we have tetradic concepts, even if we often lexicalize them with words like ‘square’ and ‘trade’, that do not take four grammatical arguments. Moreover, given a saturationist conception of composition, one might expect verb phrases that exhibit the following structure: [[[sold a horse] her] [ten dollars]]. So perhaps human I-languages don’t allow for verbs that take four arguments. And perhaps that is because phrasal combination does not signify saturation.

Similarly, we seem to have triadic concepts that get lexicalized with words like ‘triangle’ or ‘between’, which we then use in odd constructions like (18), as opposed to (18a).

(18) Sadie is between a dog and a cow.
(18a) *Sadie between a dog a cow.

But if we could label between(_, _, _) with a verb that takes three arguments, why do we circumlocute and use (18)? Studies of ditransitive constructions like (19)

(19) Hubbard gave a dog a bone.

also suggest that the constituent verbs do not combine with three arguments corresponding to saturaters of give(_, _, _); see Larson (1988), Baker (199x), and references there. Rather, it seems that (19) should be analyzed as a superficially different variant of (20),

(20) Hubbard gave a bone to a dog.

whose prepositional phrase corresponds to a monadic conjunct in a concept of the following form: \( \Phi(\)\[\text{recipient}(\_,\_\)]\(\text{dog}(\_))\).

This is the obvious analysis for (21) and (22).

(21) Hubbard kicked a bone to her dog.
(22) Hubbard kicked a dog a bone.

But if ‘give’ lexicalizes a concept that has a slot corresponding to ‘a dog’ in (19) and (20), one wants to know why ‘give’ does not express/fetch the concept \( \lambda x. \lambda y. \lambda z. T \equiv \text{give}(x'', x', x) \) — or (using ‘e’ mnemonically) \( \lambda x. \lambda y. \lambda z. \lambda e. T \equiv \text{agent}(e, x'') \& \text{chase}(e) \& \text{patient}(e, x') \& \text{instrument}(e, x) \) — and always require three saturating arguments, as suggested by the “surface form” of (19). My proposed answer is that words cannot fetch supradyadic concepts. A verb cannot fetch \( \lambda x. \lambda y. \lambda z. T \equiv \text{give}(x'', x', x) \) any more than it can fetch \( \lambda x. \lambda y. \lambda e. T \equiv \text{chase-of-by}(e, x', x) \). A human I-language allows for some words that fetch atomic dyadic concepts, subject to constraints, but not words that introduce further polyadicity. And if ‘give’ fetches a concept that has no variable corresponding to indirect objects, this already tells against the idea that lexical items have grammatical valences that reflect the adicities of lexicalized concepts.

2.3 Flexibility is Normal
If a lexical item L labels a concept that has some adicity n, then other things equal, one would expect L to combine with n arguments in any complete sentence where L does not appear as an argument of a higher-order predicate. But lexical predicates seem not to exhibit fixed selectional features concerning the number of mandatory/permitted grammatical arguments. On the contrary, it seems that a concept is often lexicalized with a word that can combine with one or two arguments—or more, if indirect objects and/or ‘that’-clauses count as arguments—to form a complete, active voice sentence. Moreover, lexical items often appear in passive forms, or nominal forms with no accompanying arguments.

---

5 I assume that ‘a horse’ is the most internal argument in (17); but this isn’t essential to the point. One wonders why ‘for’ is required in (17) if ‘ten dollars’ is an argument of the verb. Perhaps for whatever reasons, ‘sold’ can only assign three cases; though cp. ‘traded a horse for a cow’ and ‘replaced a horse with a cow’.
For example, ‘kick’ is not limited to transitive constructions like (23).

(23) Hubbard kicked a bone.
Along with (22), (24-26) are also possible.

(22) Hubbard kicked her dog a bone.
(24) A bone was kicked.
(26) Hubbard gave the bone a swift kick.

One can say that ‘kick’ is passive in (24), unaccusative in (25), and a noun in (26). But the point remains: ‘kick’ need not appear with two arguments. This flexibility bolsters independent arguments for “separating” thematic roles from lexical items, with grammatical arguments corresponding to independent conjuncts in complex concepts, and verbs corresponding to monadic concepts of events; see Schein (1993, 2002). But whatever the best diagnosis of lexical flexibility, if ‘kick’ lexicalizes a concept that has a slot for kickers, this raises the question of why (24) and (26) are even possible constructions; see Borer (2005), Ramchand (2008).

We face further questions, given the need for “event positions,” if combining verbs with arguments indicates function-application. Even if we set (25) and (26) aside for special treatment, we should ask if the meaning of ‘kick’ is triadic, with a position corresponding to the subject of (23). If the verb meaning is like $\lambda X.\lambda X'.\lambda E.\mathsf{T} \equiv \text{Kicked}(E, X', X)$ in this respect, then not only does the passive (24) call for special treatment, active voice phrases like ‘kicked the ball west’ have dyadic meanings like $\lambda X'.\lambda E.\mathsf{T} \equiv \text{Kicked}(E, X', \text{TheBall}) \& \text{West}(E)$. In which case, adding an adverbial modifier prior to the subject requires more than mere predicate conjunction: ‘west’ has to target the correct variable in a dyadic predicate; see Chung and Ladusaw (200x).

This invites the thought that verb meanings are more like $\lambda X.\lambda E.\mathsf{T} \equiv \text{Passive-Kick}(E, X)$; where the internal variable ‘X’ corresponds to internal arguments of the verb (see Kratzer 199x, Harley 200x). But then verb phrases like “kicked a ball” express monadic concepts, even if verbs themselves express concepts of certain event-participant pairs. This is already a large step in the direction urged here. Moreover, if combining a verb with an internal argument indicates saturation of a relational concept, one wants to know why verbs don’t regularly express concepts of polyadic relations that are exhibited by independent entities.

We can instead view (22-26) as evidence that verbs do not express such concepts. And however we encode event positions, we need to distinguish (23) from this sentence minus whatever aspect of it corresponds to existential closure of the ‘E’-variable. Note that (27)

(27) That Hubbard kicked it.

cannot be understood as a sentence which means that the demonstrated event was a kicking by Hubbard of a bone: while Kicked(That, Hubbard, IT) seems to be a fine thought, we cannot express it with (27). One can say that event positions of verbs are special, perhaps in ways related to tense. But overt adverbial quantification is possible, as in (28);

(28) Usually, Hubbard kicked her dog a bone after dinner.

see Lewis (1975b), Kratzer (1986, 2012). So the question is why the remaining event position in (27) cannot be overtly saturated; see §2.1 of chapter six. My answer is that combining expressions does not signify saturation. In any given phrase, the grammatical arguments of a verb V are phrasal constituents to which V bears certain structural relations, which correspond to thematic concepts that can be conjoined with others. In (26), ‘kicked’ bears the relevant

6 Perhaps several words share the pronunciation of ‘kick’. But it is no explanation to say that each such word labels a different concept. If the noun corresponds to Kick(_), one needs reasons for saying that the verbs do not.
grammatical relation to ‘That Hubbard’. But the verb cannot bear this relation to ‘Hubbard’ and another one to ‘That’; there is no third structural relation for ‘kicked’ to bear to ‘That’.

On this view, (27) presents no special difficulty. The hard questions are more about what sentences are, and how they differ from clauses like ‘Hubbard kick a bone’, if phrasal composition always yields an instruction for how to build a monadic concept. I return to this chapter six, where I show how a “complete sentence” can be viewed as an instruction for how to build a special kind of monadic concept, rather than a thought of type <t>.

2.4 What’s in a Name

Saturationist conceptions of semantic composition are initially motivated by sentences like (28), (28) Fido barked.
in which a name combines with a predicate. This is unsurprising, since Frege’s hierarchy of concepts is rooted in appeal to thoughts and denoting concepts. So if lexical items label concepts, one expects proper nouns to be parade cases, with the subject of (28) standing for a denoting concept that can saturate \( \lambda x. T = \text{Barked}(x) \) or perhaps \( \lambda x. \lambda e. T = \text{Barked}(e, x) \). But empirical and theoretical considerations suggest that lexical proper nouns do not fetch concepts of type <e>. Rather, such nouns fetch monadic concepts of a special sort, as suggested by (29) and (30), (29) Each of the Fidos barked.

(30) That Fido is the biggest one.
according to which the demonstrated Fido (as opposed to some other) is the biggest Fido.

Similar remarks apply to deictic pronouns, as in (31).

(31) He barked.
As we’ll see, the subject of (31) is not plausibly viewed as a context sensitive lexical item of type <e>. The meaning of ‘he’—as opposed to ‘she’, ‘they’, ‘it’, etc.—along with the need for indices as in (32), suggests that the subject of (31) is more like the noun phrase in (33).

(32) Fido\(_1\) barked when he\(_1\) arrived.

(33) A certain male animal barked.
And examples like (34) invite an independently plausible idea about proper nouns.

(34) That individual called ‘Fido’ is the biggest one.

These nouns, which can combine with (perhaps covert) determiners/demonstratives to form complex subjects, fetch metalinguistic monadic concepts; cp. Burge (1973), Katz (1994). That raises the question of why lexical proper nouns do not simply label concepts of type <e>. My suggestion is that human I-languages do not tolerate expressions of type <e>. In the course of lexicalization, denoting concepts are used to introduce monadic concepts that can be fetched and conjoined with others. If that’s right, human I-languages may not tolerate expressions of type <t> either. Saturation and truth may be features of thought, not the linguistic faculty that lets us generate instructions for how to build monadic concepts. But to make that suggestion plausible, I need to flesh out the argument sketched in this first section.

3. Lexical Flexibility

Suppose children lexicalize concepts that exhibit a range of Fregean types. If lexical items of a human I-language can exhibit the same range of types, and combine in Frege-Church fashion, then one would expect lexical valences to reflect the adicities of lexicalized concepts. So if this expectation turns out to be wrong, we need to consider other conceptions of human I-languages. Perhaps these biologically implemented procedures let us recursively generate (instructions for how to build) concepts of a special sort, and lexicalization lets us employ available concepts to introduce formally new and less diverse concepts that can be systematically combined in useful ways. In my view, the generalizations mentioned above point in this direction, and there is little
if any evidence for saying that lexical valence reflects conceptual adicity. To be sure, (35-38)
(35) Hubbard’s dog arrived.
(36) *Hubbard arrived a dog.
(37) Hubbard put a bone on the floor.
(38) *Hubbard put a bone.

illustrate contrasts that must be accommodated. But examples of lexical constraint do not show
that verbs have valences, at least not in any structural sense that explains why (36) and (38) are anomalous. We also need to accommodate the kind of lexical flexibility that ‘kick’ illustrates. Though before discussing words in more detail, let me enter a caveat about concepts.

3.1 Adicity Uncertainty
My view is not that speakers can intuit the adicities of lexicalized concepts. Pretheoretic
suspicions may reflect, among other things, statistics regarding how often a given verb is used
with how many arguments. Indeed, when a child acquires a word, there may not be any one concept that gets lexicalized; talk of “the” concept lexicalized may be a simplification. Some may doubt that children have concepts—mental representations that exhibit adicities—prior to having words. But I don’t see how to explain human lexical acquisition, or even animal
navigation, without ascribing concepts to creatures that do not (yet) have words. Still, one can
have good reasons for positing lexicalized concepts of diverse adicities without knowing which adicity any particular lexicalized concept has. And it is worth being clear about this, since part
of the argument for my account of semantic composition does rely on a general assumption:
children lexicalize concepts that are not already massively monadic, occasionally dyadic, and
systematically conjoinable in ways that support the cognitive productivity described in chapter
one. Again, I think lexicalization introduces concepts that can be M/Θ-joined, whatever the
adicities of the concepts lexicalized.

Recall from chapter two that Frege showed how to reformat ordinary arithmetic thoughts,
formulated in terms of concepts like NUMBER(_), in terms of logically interesting concepts like
NUMBER-OF[_, Φ()]. For purposes of logic, one can bracket the question of whether ordinary
humans somehow “have” the higher-order polyadic concepts, albeit inchoately. The important
point here is that Frege’s Begriffsschrift does not merely let us express thoughts in a way that
supports communication; it lets us re-present conceptual contents in ways that permit fruitful
definitions and new uses of certain formal operations. In my view, humans use I-languages to re-
present conceptual contents by introducing logically uninteresting (but M/Θ-joinable) concepts
like CHASE( _) and GIVE(_). So I do assume that in many cases, such concepts were not already
available for labeling via lexical items that simply pair the concepts with pronunciations.

This is not a strong claim about what was available in principle, given the expressive
power of pre-human languages of thought. An animal might have the concepts and capacities
needed to introduce CHASE(_), yet never exercise the capacities as a human child does, frequently
and spontaneously. Once my proposal is in place, there will also be room for a modification,
according to which the reformatting job that I assign to lexicalization was partly done by an older
biological innovation. Perhaps we had ancestors who used (and cousins who use) diverse

7 Though see Lidz. et.al. (201x) for interesting experiments inspired by Gordon (199x). One might have hoped to
reduce talk of KICKED(_, _) to talk of ‘kicked’ and “pre-conceptual” mental representations. But especially in light of
examples like (22-26), I see no way of avoiding appeal to at least one concept corresponding (and prior) to the word.
8 Given our natural concepts, we can represent flexibility with invented concepts that have no fixed adicity; see, e.g.,
Fine (2007). Perhaps natural concepts can likewise be “multi-variate.” But I find it hard enough to think about
concepts that have fixed adicities. And in any case, lexical items still seem to be monadic or dyadic.
concepts to introduce atomic concepts that can be joined, in ways that the initial concepts cannot, but without a capacity for pronunciation. But in any case, the cognitive productivity of human thought is mirrored by the systematic combinability of our words. And human children have a special capacity for lexical acquisition. So this capacity is presumably related to whatever is special about human thought.

I don’t deny that many lexicalized concepts are monadic. But in any particular case, it is hard to know. And this point is not limited to verbs. Perhaps ‘number’ lexicalizes a monadic concept, and Frege’s relational analog was invented. But ‘triangle’ may lexicalize a naturally triadic concept; cp. ‘between’. We have sentences like (39) rather than (40), suggesting that ‘triangle’ is used to fetch a monadic concept.

(39) That is a triangle.

(40) *This side triangles that side the hypotenuse.

Though it doesn’t follow that children simply label TRIANGLE( ) with ‘triangle’. Lexicalizers may use a polyadic concept to introduce TRIANGLE( ); where the underlining indicates an introduced concept, though perhaps one that is also available to other animals. Similar points apply to ‘square’, ‘trio’, ‘quartet’, etc. This raises familiar questions. Do words for artifacts lexicalize concepts that relate constructed things to purposes? Do some nouns lexicalize concepts that relate things to parts/histories/times/places? But for just these reasons, one might be skeptical of any generalization like “common nouns lexicalize monadic concepts.”

Recall ‘mortal’. This classical example of a monadic predicate arguably lexicalizes a concept that relates individuals to events of death. We can speak of mortals, who fall under the concept of being (a) mortal. But we can also speak of mortal wounds, which are fatal. This is not to deny that we have an atomic concept MORTAL( ). But as discussed in chapter two, atomic concepts need not be primitive; cp. MARE( ). And reflection on cases suggests that monadic concepts often mask underlying relationality. From this perspective, there was something right about the classical/medieval focus on monadic predicates, at least with regard to the distinctively human thoughts that are also natural for humans; although Frege was also right to stress the importance of a logic that accommodates relational notions. I’ll return to this point.

Likewise, I assume that many verbs are results of lexicalizing polyadic concepts, even if the details are unclear. Children presumably have a concept—EAT( _, _ ) or EAT( _, _, _ )—with which they can think about some things as (the agents of events of) eating other things. But we humans can also think about some things as (the agents of events of) grazing, dining, and “fueling up.” We also know that cows who graze “graze on” something, even though we don’t say ‘Cows graze grass’. So it isn’t crazy to think that both ‘eat’ and ‘graze’ lexicalize relational concepts yet fetch monadic concepts. Moreover, (41-44)

(41) John ate a tack.  (42) John ate something.
(43) John had a snack.  (44) John ate.

suggest that ‘eat’ lexicalizes a concept that is somehow normative; cp. ‘dine’.

Note that (41) implies (42) only on a purely existential reading of (42) that does not imply (43); and so read, (42) does not imply (44). In this sense, (42) and (44) differ in meaning. Likewise, (41) does not imply (44), unless it is assumed that a tack is nutritive for John; see Chomsky (1986). So even if ‘ate’ has a grammatical valence of -2, and (44) has a covert direct object, it doesn’t follow that the verb’s valence matches the adicity of the lexicalized concept, which may have a further slot for the relevant norm. Of course, a concept like NUTRIFY( _, _ ) or REFUEL( _, _ ) can be satisfied by event-individual pairs, and be normative in the relevant sense. But just as a “normative consumption concept” need not have a slot for a norm, it need not have
a slot for a thing consumed. So whatever one says about the grammatical valence of ‘eat’, it is a substantive hypothesis that this valence matches that of any particular lexicalizable concept. More generally, it hard to see how the number of required arguments for verbs of consumption—‘eat’, ‘dine’, ‘snack’, ‘nosh’, ‘nibble’, ‘graze’, ‘refuel’, ‘drink’, ‘gulp’, ‘sip’, ‘chug’, etc.—could generally track the adicities of the lexicalized concepts.

This is compatible with various views about how the meanings of these verbs are related. But given that ATE-A-TACK(____) is not the result of imposing a restriction on ATE(____), but more like INGESTED-A-TACK(____), one possibility is that the verb in (44) has a covert internal argument that corresponds to a conjunct like ∃[PATIENT(____)APPROPRIATE(____)]; where the normative monadic concept applies to things that are “fit” for the agent in question. Alternatively—or perhaps expressing the same idea another way—the address for ‘eat’ may be shared by a concept CONSUMED(____) that applies to events in which the thing consumed is nutritious for the relevant agent, and a more permissive concept INGESTED(____). Then the instruction fetch@‘eat’ can be executed in more than one way. But perhaps the covert argument in (44) is an instruction such that executing the verb phrase requires that ‘ate’ be executed by fetching CONSUMED(____); cp. ‘France is hexagonal’. I return to analogous issues in section four, in the context of thinking about “count to mass” implications like ‘The dragons ate a cow, so each dragon ate some beef’.

3.2 Denoting Concepts and Proper Nouns

With the caveats in mind, let’s go back to the point about names noted in section two. This can help bring out how empirically implausible the “lexicalization as labeling” idea is, at least if we assume that humans and other animals have denoting concepts. And if one wants to argue for a saturationist account of semantic composition, it does not help to deny that children have denoting concepts.

Consider (45-48), which suggest that ‘Caesar’ can be used to fetch a monadic concept.

(45) Every Caesar I saw was a tyrant.
(46) Every tyrant I saw was a Caesar.
(47) There were three Caesars at the party.
(48) That Caesar stayed late, and so did this one; but the other Caesar left early.

More precisely, there seems to be a lexical proper noun (LPN) ‘Caesar’ that is like the common noun ‘tyrant’ in this respect. Of course, names differ from common nouns. And evidently, given the contrast between (49) and (50), the subject of (49) is not a mere LPN.  

(49) Caesar arrived.
(50) *Tyrant arrived.

The subject of (49) is a name. But names may be complex expressions, consisting of an LPN and a determiner akin to ‘That’ in (48); where this determiner, covert in English, combines with LPNs but not common nouns. The sound of ‘Caesar’ might be the sound of both an LPN and a determiner phrase whose head is silent. This hypothesis is not ad hoc, given overt analogs of the posited determiner in other languages. For example, Spanish allows for both ‘Juan’ and ‘El Juan’ as devices for referring to a certain Juan. And even English allows for ‘our John’ (‘the John I know’, etc.) as a way of referring to a certain John who is suitably related to the speaker.  

---

10 Similar remarks apply to Basque, German, Scandanavian languages, many dialects of Italian, and Greek (discussed below). One can call LPNs names. But I assume that for purposes of taxonomizing human linguistic expressions, theorists need to speak of nouns and subcategories thereof; cp. Chomsky (1965). Segal (2001) offers considerations that seem to favor <e>-type analyses of proper nouns. But in my view, these considerations can be accommodated by saying that LPNs lexicalize singular concepts and fetch monadic concepts.
I won’t defend a specific proposal about the posited covert functional element. But to illustrate, and to stress that complex names can be used to designate rigidly (cp. Kripke 1982), imagine an indexable determiner $D$ such that relative to any assignment $A$ of values to indices: for each entity $e$ and index $i$, (i) $e$ is a value of $'D_i'$ iff $A$ assigns $e$ to $i$, and (ii) $e$ is a value of $'D_j$ Juan’ iff $e$ is a value of both $'D_j'$ and ‘Juan’. As discussed in chapter six, any such analysis can be recast in terms of instructions for how to build monadic concepts, and extended to pronouns: $e$ is a value of $'D_i$ she’ relative to $A$ iff $A$ assigns $e$ to $i$, and $e$ is female. One can further hypothesize that the indices index individuals via denoting concepts stored in memory, thus fostering the illusion that names are semantically atomic expressions of type $<e>$. This leaves room for various views about the meanings of LPNs. In chapter one, I pretended that an LPN fetches a partly saturated concept of identity. But given multiple Caesars, it won’t do to say that ‘Caesar’ fetches IDENTICAL(_ , CAESAR). The LPN might, however, fetch a concept that applies to individuals called—or called with the pronunciation of—‘Caesar’; cp. Katz (1994). That would accommodate the meanings of (45–48).

Like any other lexical item, the LPN ‘Caesar’ pairs a pronunciation $\pi$ with a meaning $\mu$. Suppose that $\pi$ initially connected with a denoting concept JULIUS. If a child has access to a relational concept CALLED(_, _) that applies to individual-pronunciation pairs, and to some concept $\Pi$ with which she can think about the pronunciation $\pi$, she might form a thought like CALLED(JULIUS, $\Pi$) and connect $\pi$ with a lexical meaning that is used to fetch the complex monadic concept CALLED(_, $\Pi$). Initially, the child might apply this concept to exactly one individual, Julius. This is compatible with Julius being called with many things besides $\pi$. Indeed, friends of Julius may have never used any English pronunciation to call him; cp. Kripke (198x) on whether Socrates was so-called. But more importantly, a child might think that Julius is the only bearer of the name-sound $\pi$, yet have the capacity to think otherwise by forming a thought like CALLED(ROMERO, $\Pi$). If the child is led to form and endorse such a thought, she might eventually conclude that more than one thing is called with $\pi$.

In short, an LPN can be used to lexicalize an atomic concept of type $<e>$, but used to fetch a complex (and in some sense metalinguistic) monadic concept that applies to more than one thing. On this view, it is not surprise that (51) is a possible sentence,

| (51) The Smiths are coming to dinner. |

and that it is easily heard as a claim about some people who share a surname. Moreover, as surnames remind us, even surface considerations suggest that many “proper names” are not grammatically atomic. The direct object of (52) seems to have two words as parts.

| (52) At noon, I saw Tyler Burge. |

Prima facie, ‘Tyler Burge’ is semantically related to ‘Tyler’ and ‘Burge’: a Tyler Burge is both a Tyler and a Burge. Though in a context where the only Tyler is also the only Burge, one can use (53) or (54) to talk about that one person. Examples like (55) extend this point.

| (53) I saw Tyler at noon. |
| (54) I saw Burge at noon. |
| (55) Professor Burge saw a certain Doctor Smith, but not the Doctor Smith. |

If ‘Professor’ and ‘Burge’ are expressions of types $<e, t>$ and $<e>$, then ‘Professor Burge’ denotes a truth value. One can reply that the honorific differs from the common noun ‘professor’. But that is a cost. It raises questions about how to accommodate ‘a certain Doctor Smith’ and

---

11 Or perhaps endorse a thought like IDENTICAL(CAESAR, ROMERO). Tiedke (201x) and Erickson (2013) develop different but compatible applications of this Burge/Katz-inspired suggestion; cp. Graff-Fara (201x).
‘the Doctor Smith’, with italics indicating the familiar kind of emphasis.

I return to ambiguity hypotheses. But if LPNs fetch monadic concepts, this presents a puzzle if nouns of a human I-language can be used as labels for denoting concepts like CAESAR. If lexicalizers could label denoting concepts with distinctive nouns, and thereby acquire names that can combine with a predicate of valence \( n \) to form a predicate of valence \( n-1 \), then one would expect lexicalizers to do just this and become adults for whom (45-48) and (51-52) are defective. One can imagine children who grow up to be speakers for whom ‘every Caesar’ is like the nonsensical ‘\( \forall x : c \)’ , where ‘\( c \)’ is a logical constant. But we were not such children. We acquired a language in which (45-48) and (51-52) are fine.

(45) Every Caesar I saw was a tyrant.
(46) Every tyrant I saw was a Caesar.
(47) There were three Caesars at the party.
(48) That Caesar stayed late, and so did this one; but the other Caesar left early.

So assuming that we had denoting concepts, and often used them to think about named individuals, lexicalizing these concepts was not a matter of simply labeling them with LPNs of type \(<e>\). Such lexicalization led to the acquisition of nouns like ‘Caesar’, which have the distributions they have. These proper nouns, like their common counterparts, show all the signs of being lexical items that fetch monadic concepts. They can be pluralized as in (47), or be constituents of complex demonstratives as in (48). And note that ‘one’, as it appears in (48), is ordinarily a pro-form for nouns that are not used to denote particular individuals.

These uses of proper nouns tell against the idea that LPNs fetch denoting concepts.\(^\text{12}\) But one might posit ambiguities. Perhaps speakers who use ‘Caesar’ to talk about two people—Julius and Romero—have three homophonous LPNs that are used to fetch JULIUS, ROMERO, and a complex concept like CALLED(\( c \), \( \Pi \) ). This posits both ‘saturating-LPNs’ of type \(<e>\) and ‘monadic-LPNs.’ Ambiguity hypotheses are notoriously hard to refute; see Kripke (1979b). But examples like (45-48) suggest that for each name-sound, there is a monadic LPN. And positing additional LPNs, with meanings of another type, is unattractive in several respects.

Since many human languages permit complex names, in which monadic LPNs combine with a determiner, children acquiring English can presumably formulate such an analysis of (49).

(49) Caesar arrived.

But if complex-name analyses are available to children, positing an analysis with ‘Caesar’ as a saturating LPN makes (49) strangely ambiguous: the first word might be either an instruction to fetch a monadic concept like CALLED(\( c \), \( \Pi \) ) or an instruction to fetch any of several denoting concepts; yet the posited denoting meanings for (49) could be plausibly redescribed in terms of the monadic meaning. In this sense, appeals to “saturating-LPNs” are not just theoretically redundant, they require children who posit lexical ambiguities beyond necessity. And one might think that children would, if possible, adopt uniform analyses of LPNs.

Relatively, appeal to saturating LPNs makes “noun” a disjunctive grammatical category, and not just because of the two semantic types. Saturating LPNs would not head phrases, much less phrases of their own type. Since common nouns can combine with other expressions to form noun phrases like ‘angry tyrant who died’, much as verbs can combine with other expressions to form verb phrases like ‘quickly ate his lunch’, one needs special reasons for saying that ‘Caesar’ is ambiguous as between (i) a monadic-LPN that can head the noun phrase ‘angry Caesar who arrived yesterday’, which can combine with a determiner like ‘every’, and (ii) a saturating-LPN.

\(^{12}\) If ‘Tyler Burge’ is as semantically distinct from ‘Tyler’ as ‘Twain’ is from ‘Clemens’, then we also want to know why ‘Tyler Burge is a philosopher.’ seems to follow from (52) and ‘Every Tyler I saw is a philosopher.’
that cannot head such a phrase, but must always appear as a complement of some predicative phrase whose valence is then different from that of its head. It is worth remembering that given a valence/saturationist conception of composition, predicate-argument phrases differ crucially from their heads. But even accepting this departure from a minimalist (“bare phrase structure”) ideal, appeal to (ii) seems to complicate grammar needlessly, given appeal to (i).

In English, (45–48) are in some sense “marked” constructions, compared with (49). This makes it tempting to think that the question is whether or not English happens to be a language in which name-sounds are ambiguous, as if both options are available to children. But it is an empirical hypothesis that saturating-LPNs are even possible human linguistic expressions. And if lexicalization can introduce monadic concepts, any appeal to saturating-LPNs stands in need of support—especially since there are human languages in which complex names, with monadic LPNs as constituents, appear regularly. Such languages highlight the question of whether or not the human language faculty even permits saturating-LPNs.

For example, in Greek, analogs of (49) are as anomalous as analogs of (50).

(50) *Tyrant arrived.

A bare LPN like ‘Petros’ is no better, at least discourse initially, than a bare common noun as the subject of a sentence. To talk about the contextually relevant person called ‘Petros’, one uses the masculine determiner ‘o’ to form the phrase ‘o Petros’. Any child can acquire such a language. And if English has saturating LPNs, along with monadic LPNs, any child can acquire such a language. So if the ambiguity hypothesis for English is correct: experience with English leads every normal acquirer to a lexicon with enough LPN entries, despite homophony and the grammatical possibility of monadic LPN analyses that would shorten the lexicon; and experience with Greek leads every normal acquirer to a lexicon without too many entries, despite the possibility of ambiguity and saturating LPN analyses that would lengthen the lexicon.

Usually, children treat lexical sounds as ambiguous only given reason to do so. So we have to ask what would lead children to conclude that English name sounds are ambiguous? One might speculate that not hearing the determiner lets children know that English has lexical names. But even if children are shielded from phrases like ‘the John we met yesterday’, they may not use “negative” evidence in acquisition; see, e.g., Crain and Pietroski (2001). Worse, a special lexical type must be posited to let children use negative evidence to acquire a grammar that admits theoretically superfluous ambiguities. By contrast, if LPNs fetch monadic concepts, languages that regularly combine LPNs with determiners are expected. English can be viewed as a slightly quieter language in which the determiners are often though not always covert.

Thus, several considerations converge to suggest that there are no saturating LPNs, even though children plausibly lexicalize many denoting concepts with LPNs. This is unexpected if human I-languages permit proper nouns of type <c>, with verb-name combinations signifying saturation of predicative concepts by denoting concepts. By contrast, if semantic composition is fundamentally conjunctive, then given ways of using denoting concepts to introduce monadic concepts, it is no surprise that lexicalizing denoting concepts involves such introduction. With this in mind, let’s return to cases in which the lexicalized concepts are plausibly polyadic.

3.3 Supradyadic Concepts and Verbs

If there are no 17-place concepts to lexicalize, then the absence of verbs with valence –17 tells us

---

13 See Giannakidou and Stavrou (1999) for discussion in a broader context of Greek nominals. There is a vocative case form, used in address, that precludes determiners (and the final ‘s’ of masculine nouns). But vocatives are plausibly used to call people, rather than fetch concepts of those called with certain LPNs. And even English has names like ‘the Thames’, ‘the Bronx’, ‘the Met’, ‘the Mets’, etc.
little. But as Gallistel (1990, 200x) discusses, it seems that animal navigation already requires polyadic representations: a forager bee can somehow relate itself to a target via the position of the sun and the time of day; an ant can somehow keep track how many steps are needed, along which vector, to relative to the hive. Humans, who can distinguish selling from giving, certainly seem to have concepts that are at least tetradic. As noted in section one, our concepts of selling and giving seem to differ, in the that former has a greater adicity: a seller gets something back in exchange for the thing transferred to the recipient. So if ‘give’ is alleged to have valence –3 because it lexicalized a triadic concept, one wants to know why ‘sell’ does not have valence -4?

We could invent a language in which (56) is a sentence with the meaning of (56a).

(56) *The colonel sold the knife the professor ten dollars.
(56a) The colonel sold the knife to the professor for ten dollars.

But in English, (56) is anomalous. And prima facie, ‘sold’ takes two arguments in (57).

(57) The colonel sold the knife.

Positing two further covert arguments seems ad hoc, especially as part of a proposal that eschews a covert constituent of names in English. Similar remarks apply to ‘bought’; though (58)

(58) Professor Plum bought Miss Scarlet the knife.

is roughly synonymous with (58a), which has a “benefactive” implication, unlike (59).

(58a) Professor Plum bought the knife for Miss Scarlet.
(59) Professor Plum bought the knife for ten dollars.

Presumably, the valence of ‘bought the knife’ is not reduced by ‘for Miss Scarlet’ in (58a), else ‘bought’ would have valence -5 given (60).

(60) Plum bought the knife from Mustard for Scarlet for ten dollars.

But then the valence of ‘bought’ is not reduced by ‘Miss Scarlet’ in (58). This bears on (61),

(61) Plum gave Scarlet the knife.

since (58) has a closely related meaning which adds the implication that Plum paid for the gift. And famously, (62) is odd while (62a) is not.

(62) *Scarlet donated the museum the art.
(62a) Scarlet donated the art to the museum.

All of which suggests that ‘gave’ and ‘donate’ do not fetch concepts that are saturated by correlates of three grammatical arguments. It seems to be an idiosyncratic fact that ‘donate’, unlike ‘give’, resists ditransitive constructions. But however the oddity of (62) is explained, the acceptability of (63-65) is puzzling if ‘give’ has valence -3.

(63) Scarlet gave the money away.
(64) Plum gave at the office.
(65) The money was given to charity.

Repeating a point from section one, the mere acceptability of (61) cannot show that ‘give’ has valence -3, else (22) would show that ‘kick’ has the same valence.

(22) Hubbard kicked her dog a bone.

And (22) is not an isolated case of “ditransitivization.” We usually think of ‘cooked’ and ‘sang’ as combining with one or perhaps two arguments, as in (66) or (67).

(66) Mrs. White cooked, while Colonel Mustard sang.
(67) White cooked an egg, while Mustard sang a lullaby.

But consider (68), synonymous with (68a), and (69).

(68) White cooked Mustard an egg, while he sang the baby a lullaby.
(68a) White cooked an egg for Mustard, while he sang a lullaby to the baby.
(69) White bounced/fed/stole/left Mustard a tennis ball.
It seems that many (intuitively transitive) verbs can appear in ditransitive constructions, in which a “recipient” is specified, without lexicalizing concepts that have a slot for recipients.

We also seem to have triadic concepts like \textsc{between}(_, _, _) that don’t get expressed with verbs that take three arguments. The claim made with (70)

(70) Plum was between Scarlet and White.

is correct only if Plum, Scarlet, and White exhibit a certain triadic relation. But (70) is a subject-predicate sentence whose predicate employs three functional words to accommodate two names. This raises the question of why we don’t have a verb ‘twixt’—of type <e, e, et>—such that (71)

(71) Plum twixted Scarlet White.

has the meaning of (70). But there is a puzzle here only if verbs can fetch triadic concepts.\footnote{Note that we can impose a meaning, Jabberwocky style, on (71): Plum did something to a certain Scarlet White; or perhaps Plum did something that involved twirling Scarlet to White.}

In this context, it is relevant that ‘jimmied’ takes two arguments.

(72) Mister Green jimmied the lock (with a screwdriver).

(72a) *Mister Green jimmied the lock a screwdriver.

Reference to an implement appears within a grammatically optional adjunct, as in (72), not as a mandatory third argument.\footnote{I am indebted to Alexander Williams for this example, and more importantly, for a series of conversations that deeply influenced the development of this section and the next. He has discussed closely related matters in Williams (2005, 2007). Note that if ‘He jimmied me the lock’ has a meaning, it is that he jimmied the lock for me, as opposed to he jimmied the lock with me, and likewise for ‘He jimmied the screwdriver the lock’. Cp. note 7.} The concept lexicalized presumably has a slot for that with which the jimmier jimmies the jimmied. So we need some explanation why (72a) is not a perfectly good ditransitive construction with the meaning of (72). Similar points apply to (73/73a).

(73) Green nailed Plum (with a devastating objection/right hook/fastball/lawsuit)

(73a) *Green nailed Plum a lawsuit

One can hypothesize that “supratransitive” verbs are disallowed for reasons stemming from the underlying syntax of human languages; see Hale and Keyser (1993), etc. But this idea, fully compatible with a “conjunctivist” conception of semantic composition, fits ill with the idea that combining expressions often signifies saturation. If a mode of composition can be employed twice in a sentence, why not thrice or more? And if syntax precludes saturating a predicate more than twice—so that at least in this sense, saturation is not available as a recursive mode of composition—we have to ask whether this mode of conceptual composition is invoked by any mode of linguistic combination.

As a final illustration of argument flexibility, consider the concept of marriage. Whatever its adicity, this concept can be indicated with a noun. Yet each of (74-78) might be used to describe the same wedding.

(74) Scarlet married Plum, but their marriage was doomed.

(75) Scarlet got married to Plum, with the Reverend Green officiating.

(76) With reservations, Green married Plum and Scarlet.

(77) Plum and Scarlet married, and they got married in a hurry.

(78) It was Scarlet’s first marriage, though Plum married for third time.

This suggests that given three acting participants—Scarlet, Plum, and Green—we can describe various events of marrying that fall under a monadic concept (of marriage) that a competent speaker has given the word ‘marry’, which lexicalizes a relational concept.

3.4 Dyadic Concepts

One might concede that when a denoting or supradyadic concept is lexicalized, the result is a
word that has a valence of −1 or −2. But this doesn't yet show that all lexical items fetch monadic or dyadic concepts. In particular, one might think that given the “event positions,” lexicalizing a relational concept often results in a word that fetches a triadic concept.

For these purposes, let’s distinguish “robustly” dyadic concepts like kick(_, _) and chase(_, _)—concepts of genuine relations exhibited by pairs of independent entities—from concepts like agent(_, _) that are formally dyadic, but apply to pairs consisting of an entity and something in which that entity participates. We can think of kick-by-of(_, _, _) as formally triadic but robustly dyadic, since it applies to <ε, α, β> iff ε is an event of α kicking β. If the concept lexicalized with ‘kick’ robustly dyadic, one might expect ‘kick’ to fetch a robustly dyadic concept, at least when it appears as an active voice verb. Similarly, if the valence of ‘give’ is capped at −2 because syntax allows only two grammatical arguments, one might expect ‘give’ to fetch a robustly dyadic concept. So if verbs that lexicalize robustly relational concepts do not fetch such concepts, one wants to know why.

I stress this point because many considerations press in the direction of saying that transitive constructions do not include verbs that have robustly relational meanings. To be sure, many examples are compatible with the hypothesis that some verbs have such meanings. But in my view, little if any evidence favors this hypothesis, once we take its negation seriously.

Given (23), it might seem that the lexical item ‘kick’ has a robustly dyadic meaning,

(23) Hubbard kicked a bone.

specified in terms of kick(_, _) or the formally triadic concept kick-by-of(_, _, _). But then (22)

(22) Hubbard kicked her dog a bone.

calls for a theoretical choice. Perhaps the verb in a ditransitive construction can have a robustly dyadic meaning, and the “indirect object” of (22) reflects a covert preposition. Or maybe the lexical item ‘kick’ is used, perhaps along with a covert functional item, to introduce a related word that has a robustly triadic meaning. In my view, each of these options is coherent. But I assume that no process can change the adicity of a concept or the valence of any particular linguistic expression.

Given the requisite conceptual tools, kick-by-of(_, _, _) and recipient(_, _) can be used to introduce a formally tetradic though robustly triadic concept kick-by-of-to(_, _, _, _) that applies to <ε, α, β, γ> iff ε is an event of α kicking β to γ. But kick-by-of(_, _, _) remains formally triadic. Likewise, one can posit a process—perhaps an ordinary process of grammatical composition—in which a verb that (by itself) takes exactly two grammatical arguments is used to create an expression that takes exactly three grammatical arguments. If ‘kick’ has a valence of −2, perhaps it can be used to create an analytically related and homophonous expression that has a valence of −3. But in any case, composition is not alchemy. If a hypothesized notion of grammatical valence is to explain why a verb can and must take a certain number of arguments (in a complete sentence), then valence cannot be a changeable property of the verb. If a verb V has a valence of −2, and a verb V’ appears (in a complete sentence) with one or three arguments, then V’ is not identically with V. It may be that one verb is a constructed variant of the other. But another possibility is that few if any lexical items have valences, in any explanatory sense, even if verbs often impose requirements on the sentences in which they appear.

16 Though it can be hard to distinguish [(kicked (ta her dog)] (a bone)] from [(kicked+t+ta (her dog)] (a bone)]

17 In terms of the 19th century chemical-compounding analogy, one might think of iron, which permits both FeO and Fe₂O₃, though if one thinks of O has having a valence of −2, one will say that Fe can have a valence of +2 or +3. So perhaps the analog of kicked+ta is a pretended ion that has a valence −3 but remains a variant of Oxygen (as opposed to, say, Nitrogen). Or perhaps the analogy should be dropped, given modern conceptions of electron shells.
I’ll return to this last point. But note that if the meaning of ‘kick’ is robustly dyadic, then (22) is hardly the only construction that calls for hard theoretical choices. Recall (24-26).

(24) A bone was kicked.
(26) Hubbard gave the bone a swift kick.

One can say that ‘kick’ permits passivization as in (24), object dropping as in (25), and nominalization as in (26). Indeed, this kind of descriptive vocabulary can be useful, at least initially: by characterizing some constructions as transformation or variants of others, one can highlight commonalities across constructions. But prima facie, (24-26) are grammatically complete expressions in which ‘kick’ does not take two arguments. I can’t prove that (25) lacks a covert direct object that does not induce the implication that the baby kicked something; nor can I prove that (25) lacks a covert indirect object that does not imply a recipient. But even if (25) has a covert direct object, as Hale and Keyser (1993) suggest, this is compatible with ‘kick’ fetching the monadic concept KICK(_). And since (25) does not imply that the baby kicked anything, theoretical considerations that motivate appeals to a covert grammatical argument do not motivate appeals to verb meanings that are robustly dyadic yet compatible with semantically anemic direct objects.

A competent speaker will know that if (24) is used to make a true claim, then there was an agent of the relevant bone-kicking. But such a speaker will also know that the bone-kicking occurred in a particular place at a particular time, that the agent had legs (or suitable analogs), etc. Competent speakers may well know that kickings have agents, especially if KICK(_) is introduced in the course of lexicalizing a robustly dyadic concept like KICK(_,__). But it doesn’t follow that (24) and (26), generable expressions of a human I-language, have constituents that somehow represent kickers. Endlessly many properties of events, essential and contingent, can go unrepresented by (instructions for how to build) concepts that can be used to think about events.

With enough ingenuity, one can preserve the idea that ‘kick’ is fundamentally a label for a robustly dyadic concept, even given constructions like (24-26). But once one appeals to processes that use certain representations to introduce other representations that have different valence properties, we face the question of whether lexicalization is itself such a process. Instead of saying that children acquire a robustly dyadic verb that can be used to introduce a passive variant that does not require a grammatical subject, one can say that when children lexicalize a robustly dyadic concept, they introduce a monadic (passive-friendly) concept and acquire a word that does not have a valence of –2. From this perspective, the main theoretical questions concern human linguistic capacities, not how we choose to describe the constructions generated via those capacities. So it is worth noting the absence of logically possible words that would be expected given a labeling conception of lexicalization.

In discussing ditransitive constructions, I noted that the absence of certain verbs—and the need for certain kinds of circumlocution—bolsters the considerations in favor of denying that supradyadic concepts are lexicalized with grammatical predicates of matching valence. Similar points apply with regard to transitive constructions and dyadic verbs. It seems that prepositions often lexicalize robustly dyadic concepts like FROM(_,_) or perhaps formally triadic analogs that permit temporal modification. So it is interesting that in (79),

(79) Plum is from Devon.

the prepositional phrase combines with a copula to form a tensed monadic predicate. One can

---

imagine a language with a corresponding transitive verb, as in (79a); cp. (80).

(79a) *Plum froms Devon.
(80) Plum hails/comes/arrived from Devon.
But it seems that ‘from’ cannot be used to label a concept like \( \lambda y. \lambda x. T \equiv \text{ISFROM}(x, y) \) or \( \lambda y. \lambda x. \lambda e. T \equiv \text{COMESFROM}(e, x, y) \). Circumlocution is required.\(^{19}\)

Similarly, we use (81), not (81a),

(81) Plum is taller than Green
(81a) *Plum talls Green

as if relational concepts cannot be lexicalized with open-class monomorphic expressions. So perhaps robustly dyadic concepts like \( \text{FROM}(\_ , \_) \) and \( \text{TALLER}(\_ , \_) \) cannot be labeled/fetched with verbs, not even verbs that take two grammatical arguments.\(^{20}\) Once we take this hypothesis seriously, standard accounts of causatives can be seen as more instances of the general trend.

It is a familiar idea that (82) and (83) have the indicated grammatical structures;

(82) Green broke the glass.  (83) The glass broke.
(82a) [Green \( [\nu \text{ broke}] \ [.i \ [\text{the glass}]]] \)  (83a) [[the glass]i \ \ldots \ [\text{broke} \_i]]

where the lexical item ‘broke’ appears as a constituent of both sentences. On this view, (82) is formed by combining the overt verb with a functional item \( \nu \), which is covert in English but overt in many other languages.\(^{21}\) It is independently plausible that the internal argument of ‘broke’ in (83) displaces to a matrix subject position without becoming the external argument of ‘broke’; see Burzio (19xx, etc). So it seems that lexical items like ‘break’, which can appear in transitive constructions, do not fetch concepts like \( \lambda y. \lambda x. \lambda e. T \equiv \text{BREAK-BY-OF}(e, x, y) \). One can say that the concept lexicalized with ‘break’ is—perhaps modulo an event variable—a monadic concept of broken things, rather than a robustly dyadic concept. But this hypothesis, about how (all?) children acquire the word ‘break’, seems rather implausible for the full range of examples.

Recall (12) and (13). Does ‘drop’ lexicalize a concept that is not robustly relational?

(12) Trees dropped leaves onto the ground, as a student dropped her books.
(13) Leaves dropped onto the ground.

Perhaps we understand ‘kill’ (Old English, ‘cwell’) as the transitive form of ‘die’ (Old English, ‘cwell’); see Traugott (1972). And perhaps we understand the verb in ‘They felled trees.’ as a transitive form of ‘fall’. It doesn’t follow that we have no concept of killing/felling that does not include a concept of dying/falling. But even if many causative verbs lexicalize complex concepts, pace Fodor and Lepore (2002), I see no reason to deny that many verbs lexicalize atomic concepts that are robustly relational.\(^{22}\) Such verbs may, however, fetch monadic concepts.

---

\(^{19}\) My own view is not that prepositions let us introduce and fetch monadic concepts of frommings. Perhaps ‘from’ can fetch a formally but not robustly dyadic concept \( \text{FROM}(\_ , \_) \) that applies to \( <\epsilon, \beta> \) iff \( \epsilon \) is an event/state of \( \beta \) coming/being from somewhere; see Schein (2002), Svenonius (2010). But given polysemy, it may well be that for any given preposition, more than one concept can be fetched from its lexical address.

\(^{20}\) One might reply that ‘tall’ is of this type, with relation holding between entities like Plum and abstracta like heights; cp. Kennedy (1999). But while I agree that ‘tall’ is used to signify a relation, in a way that invites talk of heights (or degrees thereof), circumlocution is still required, as Kennedy’s work highlights.

\(^{21}\) See, e.g., Chomsky (1995) and Kratzer (1996), drawing on Baker’s (1988) revival of an old idea. For further references and discussion in a neo-Davidsonian context, see Pietroski (2003b).

\(^{22}\) Given how little we know about lexicalized concepts, ‘break’ may be an example. A child might use \( \text{BREAK}(x, y) \) or \( \text{BREAK}(e, x, y) \) to introduce both \( \text{BREAK}(e) \) and another monadic concept \( \text{BREAK}(f) \) such that: \( \exists [\text{BREAK}(f) \& \text{TERMINATE}(e, f)] \equiv \text{BREAK}(e) \); where \( \text{BREAK}(f) \) applies to events of something breaking, regardless of the cause, and \( \text{TERMINATE}(e, f) \) is a concept of a whole-to-part relation that processes bear to their final parts. Cp. Levin and
In any case, once causative constructions like (77) are set aside, we need to ask how many of transitive constructions plausibly include a lexical verb that has a valence of –2. Since (25) does not imply that the baby kicked, one have hoped that (23) would be a paradigm case.

(23) Hubbard kicked a bone.

I think the following quote—from exceptionally good psycholinguists, who have thought long and hard about lexicalization—reflects a common but mistaken view.

Clearly, the number of noun phrases required for the grammaticality of a verb in a sentence is a function of the number of participants logically implied by the verb meaning. It takes only one to sneeze, and therefore *sneeze* is intransitive, but it takes two for a kicking act (kicker and kickee), and hence *kick* is transitive (Lidz, Gleitman, and Gleitman 200x).

The authors immediately qualify this remark, saying that “there are quirks and provisos to these systematic form-to-meaning-correspondences.” But even setting (22) and (26) aside as quirks,

(22) Hubbard kicked her dog a bone.
(26) Hubbard gave the bone a swift kick.

one might wonder if (25) and (24) are plausibly accommodated with provisos.

(24) A bone was kicked.

Prima facie, (25) challenges the idea that an *act* of kicking requires two participants, even if one grants that paradigmatic acts/processes of kicking do involve both a kicker and a kickee. The important truth behind the quote is that the paradigmatic cases, when described with transitive constructions like (23), may well be the ones that help children figure out which concept to lexicalize with ‘kick’. And this may feed our sense, as adults, that the transitive uses of ‘kick’ are somehow primary. But in my view, passive constructions like (24) are neither grammatical quirks nor by-products of a special process that can somehow make a transitive verb intransitive. We should, I think, be puzzled by talk of passivization (and nominalization). If the meaning of ‘kick’ somehow involves a variable position for kickers—and this explains why ‘kicked a bone’ is a mere verb phrase as opposed to a “semantically complete” sentence—then how can a word have the meaning of ‘kick’ but no argument position for kickers? I think that ‘kick’ can appear as a passive verb (or even as a noun), despite lexicalizing a *concept* that has an unsaturated slot for kickers, because lexical items need not fetch the concepts they lexicalize. Perhaps there are other explanations for why (24) and (25) can be understood as complete sentences. But I don’t think the facts are *explained* by saying that verbs have valence properties whose canonical manifestations are limited to active voice uses.

If only because passive uses seem to manifest core grammatical competence, one might follow Kratzer (1996) and “sever” variables for agents from verb meanings. One can say that the lexical item ‘kick’ fetches $\lambda Y.\lambda E.T \equiv \text{PASSIVE-KICK}(E, Y)$; where this concept is perhaps introduced in terms of the robustly dyadic $\lambda Y.\lambda E.\lambda X.T \equiv \text{KICK}(E, X, Y)$. On this view, transitive uses of ‘kick’ require a second “small” verb that supports addition of an external argument, as indicated in (23a); cp. Chomsky (1995).

(23a) Hubbard $[v \text{[kicked a bone]}]\]

For these purposes, we needn’t worry about whether the verbs “incorporate;” cp. Baker (1988). The idea is that one way or another, [kicked a bone] and [Hubbard $[v \text{[kicked a bone]}]\]

---

Rappaport (1995). See Pietroski (1998, 2003b, 2005a) for discussion of TERMINATER($E, X$), its relation to PATIENT($E, X$), and the extensive literature on these topics.
correspond respectively to the distinct monadic concepts \( \lambda e. T = \exists y[\text{PASSIVE-KICK} & \text{BONE}(y)] \) and \( \lambda e. T = \text{AGENT}(e, \text{HUBBARD}) & \exists y[\text{PASSIVE-KICK} & \text{BONE}(y)]. \)

One can preserve the idea that arguments of verbs saturate verb meanings, by assigning to \( \nu \) a meaning of type \( \langle e, t, e, t \rangle \) such that \([\nu \text{ [kicked a bone]}\) corresponds to the formally (but not robustly) dyadic concept \( \lambda x. \lambda e. T = \text{AGENT}(e, x) & \exists y[\text{PASSIVE-KICK} & \text{BONE}(y)]. \) But as noted above, my view permits a variant according to which ‘kick’ fetches the introduced concept \( \text{PASSIVE-KICK}(\_, \_); \) where this concept applies to certain event-participant pairs. Then one can say that [kicked a bone] and \([\nu \text{ [kicked a bone]}\) both correspond to \( \exists[\text{KICK}(\_, \_) \text{Bone}(\_)]. \) One needn’t assign a special meaning to \( \nu \) itself in order to say that being an external argument of a verb has thematic significance. And if \( \nu \) mediates a thematically significant relation that ‘HUBBARD’ (but not ‘a bone’) bears to ‘kicked’ in (23a), then we can treat (23a) as an instruction for how to build a concept like \( \exists[\text{AGENT}(\_, \_) \text{THAT-HUBBARD}(\_)] \exists[\text{KICK}(\_, \_) \text{Bone}(\_)]. \) But compositional details aside, I think Kratzer’s proposal is an important step in the right direction. Once variables for agents are severed from verb meanings, we need to rethink standard conceptions of how lexical meanings are related to concepts lexicalized.

In this context, recall that (84) lacks the reading indicated with (84c); see chapter zero.

(84) a spy saw a man with a telescope
(84a) A spy saw a man using a telescope.
(84b) A spy saw a man who had a telescope.
(84c) #A spy saw a man and had a telescope.

The prepositional phrase can be understood as modifying ‘saw a man’ or ‘man’. But in the former case, the implication is that an event of seeing was done with a telescope, and not merely that a spy was in possession of a telescope. Similar remarks apply to (85).

(85) A doctor rode a horse from Texas.

If ‘see’ and ‘ride’ have robustly relational meanings like \( \lambda y. \lambda e. \lambda x. T = \text{SEE-BY-OF}(e, x, y) \) and \( \lambda y. \lambda e. \lambda x. T = \text{RIDE-BY-OF}(e, x, y) \)—with variables corresponding to the grammatical subject—the question is why we can and evidently must skip over this variable when interpreting the prepositional phrase. Like Kratzer, I think the verb meaning provides no such variable to skip over. One can instead posit a special rule of composition that encodes how we can understand (79); given two expressions that have meanings of the form \( \lambda e' \lambda e. T = R(e, e') \) and \( \lambda e. T = M(\_); \) combining the expressions creates a phrase with a meaning of the form \( \lambda e' \lambda e. T = R(e, e') \& M(e); \) see Chung and Ladusaw (2003). But this way of describing the facts does not explain them, absent an account of why the phrasal meaning is not of the form \( \lambda e'. \lambda e. T = R(e, e') \& M(e'). \)

If there were no independent arguments for “thematic separation,” and lots of evidence that supports a saturationist conception of semantic composition for verb phrases, it might be reasonable to set this issue aside for future work. But at some point, we have to ask what data actually confirms (as opposed to being consistent with) the claim that verbs can and often do have robustly relational meanings. If ditransitives and causatives and passives are special cases, one wonders which “core cases” motivate treating the others as somehow quirky. Reports like (86-88) are often set aside as special cases. But they may be instructive.

(86) It is snowing in London, where it often rains.
(87) Snow is falling, but at least it isn’t raining frogs.

---

23 If it helps, suppose that in the imagined language, this reading can be expressed with ‘From Texas, a doctor rode a horse’. One might speculate that in English ‘from Texas’ cannot combine with ‘ride the horse’, but must rather attach “higher.” But this seems unmotivated, at least as a strategy, given the parallel point about ‘with binoculars’. 
A spy saw it snow in London as we watched rain fall in Venice. Prima facie, the verbs ‘snow’ and ‘rain’ fetch event-concepts that lack a variable for that which falls when it snows/rains. But one can use a direct object to say what is falling; and if ‘rained frogs’ and ‘rained blood’ are grammatical, along with ‘rained dihydrogen oxide’, it is hard to see how ‘rained water’ could be ungrammatical as opposed to redundantly weird.

From a saturationist perspective, (86-88) are puzzling. Given the need for event variables, the verbs in ‘It rained/snowed/poured/drizzled’ cannot be treated as devices for expressing thoughts with no unsaturated elements, even if there are such thoughts; cp. Montague (1974). But if ‘rain’ expresses RAIN(E), and the argumentless verb corresponds to an argumentless concept modulo the event variable, we need some explanation for why (89) is acceptable and why it implies (90).

(89) Rocks rained down on the village.
(90) Rocks fell on the village.

An obvious initial suggestion is that RAIN(E) is an essentially plural variant of FALL(E), which is introduced via FALL(E, X), a concept that relates falls to fallen; cp. Boolos (1998). If some events satisfy RAIN(E), they are falls; if their patients were rocks that ended up on the village, they were falls of rocks that ended up on the village. And if we typically use ‘rain’ to think/talk about waterdrops, we might add a nominal use as in (88).24 My aim is not, however, to provide a theory of weather reports. It is rather to highlight two points. First, if verbs express monadic concepts of things that can have participants, then “argument optionality” is not surprising. If ‘rain’ expresses RAIN(E), then absent lexical restrictions of the sort imposed by ‘put’, (86-90) do not present puzzles.

3.5 LANs, SCANs, and POSSEs
In thinking about the difficulties for any view that eschews appeal to verbal valences, it is useful to distinguish cases in which certain arguments seem to be required from cases in which certain arguments see to be prohibited. As illustrations of the first type, consider (91) and (92).

(91) *John devoured.
(92) *John put.
(91a) #John devoured something.
(92a) #John put the book somewhere.

As indicated, not only are the shorter sentences anomalous, they are not understood as slightly defective ways of pronouncing the meanings of (91a) and (92a); cp. ‘seems sleeping/sleepy’.

As illustrations of the second type, consider (93) and (94), which are quite defective,

(93) *Brutus sneezed Caesar.
(94) *Brutus arrived Caesar.
(93a) Brutus made Caesar sneeze.
(94a) Brutus caused Caesar’s arrival.

but can be heard as having the meanings of (93a) and (94a); cp. ‘burped the baby’.

And of course, these are not the only facts.

Recall that (42) does not follow from (44), which does follow from (95),

(42) John ate something.
(44) John ate.
(95) John dined.

24 Note that ‘Cats and dogs rained down on Rome’ does not have the idiomatic meaning of ‘It rained cats and dogs in Rome’, which is roughly that it rained heavily in Rome. One might argue that ‘snow’ expresses SNOW(E, L), with a variable for locations. But even if this is correct, it is little comfort to saturationists. For unlike the variable for the fallen in FALL(E, X), the location variable is not saturated by the concept expressed with any argument of the verb. We can say ‘Snow fell’ and ‘Rome fell’, but not ‘Rome snowed’. And if one insists that ‘It snowed’ has a covert saturating location argument, as opposed to a covert conjoining location adjunct, one needs appropriate analyses of (89) and (90).
even though the preposition in (96) is obligatory,

(96) John dined on shrimp.

as if the concept expressed with ‘dine’ does not have a variable for the food eaten—even though ‘dine’ imposes some condition on the manner of eating and/or food eaten. Any theory will have to encode, one way or another, the differences among ‘ate’, ‘dine’, and ‘devour’. One option is to say that ‘devour’ always has a valence –2, while ‘ate’ can appear in either (a primary) transitive form that has valence –1 or an (object-dropping) intransitive form that has valence –1. But in my view, this sacrifices the potential explanatory force of saying that verbal valences are determined by lexicalized concepts. And as a description of the facts, I don’t see why the two alleged forms of ‘ate’ should be viewed as “variants” rather than homophones, perhaps linked by a meaning postulate that captures the (one-way) implication from (44) to (42).

We do need to account for the anomaly of (91), since there seems to be a perfectly fine corresponding thought: John was the agent of a devouring. If (44) is an instruction for how to build a thought like ∃[AGENT(_ , _)^JOHN]^ATE(_), then other things equal, (91) should be an instruction for how to build a thought like ∃[AGENT(_ , _)^JOHN]^DEVOURED(_). So if ‘ate’ does not require overt specification of the eaten, why does ‘devour’ require overt specification of the devoured? This is a fine question. But the contrast between (91) and (44) does not support the idea that verbs have valences. If we diagnose the anomaly of (91) in terms of a “missing” internal argument, then the puzzle is why (44) is acceptable. One can say that ‘ate’ has a covert direct object in (44); but that seems less plausible for ‘dined’ (‘snacked’, ‘noshed’, etc.). Moreover, if (44) includes a covert object that satisfies the demands of ‘ate’, why can’t (91) be understood as including a covert object that satisfies the demands of “devoured”?

If appealing to valences was required to account for the anomaly of (91), we might have to live with the “two forms” account of why both (42) and (44) are acceptable. But here, ‘put’ is instructive, and in some ways a helpful example for me. Consider (97), which is also anomalous.

(97) *Chris put the book.

Appeal to valences accounts for the unacceptability here only if ‘put’ has a valence of -3. This is the prediction if verbs inherit their valences from concepts lexicalized, and ‘put’ lexicalized a robustly triadic concept (with a slot for where the agent puts the thing in question). But in (92a), ‘somewhere’ seems to be an adjunct rather than a grammatical argument of the verb. So to maintain a valence explanation for why (97) is unacceptable, the notion of valence must be understood so that a verb with a valence of –n can combine with fewer than n arguments to form a grammatically complete active voice construction—at least if n exceeds 2—so long as the “valence remainder” corresponds to a (non-optional) adjunct. One can make this hypothesis seem less ad hoc by claiming that human I-languages permit at most two arguments per verb. I think this plausible claim about grammar fits ill with the idea that grammatical arguments correspond to saturable slots of concepts fetched by verbs. But in any case, if a verb can require an “associate” that is not an argument, why not say that a verb can require an associate that is not a valence-reducer?

We can say that ‘put’ demands a location-specifier, while neither ‘eat’ nor ‘devour’ do, and that ‘devour’ demands a theme-specifier in a way that ‘eat’ does not. In the latter respect, devour is like ‘fell’ and unlike ‘rain’. One way or another, such lexical idiosyncracies must be encoded. Some verbs are more flexible than others with regard to how many arguments they can/must appear with in active voice constructions. But that is not yet a reason for encoding the idiosyncracies in terms of valences and the idea that a flexible verb has more than one form.

Let’s say that each lexical item L has a “Semantic Composition Adicity Number”
(SCAN) corresponding to the adicity of the concept(s) fetched with L. A lexical item that always fetches a monadic concept has a SCAN of −1; a lexical item that always fetches a dyadic concept as a SCAN of −2.²⁵ Let the “Lexicalized Adicity Number” (LAN) of L reflect the adicity of the concept lexicalized with L: if L lexicalized a dyadic concept, its LAN is −2; if L lexicalized a triadic concept, its LAN is −3; etc. Finally, let’s say that each verb also has a “Property of Smallest Sentential Entourage” (POSSE), corresponding to the number of grammatical arguments and/or mandatory adjuncts that must appear with the verb in a grammatical active voice sentence. Given this terminology, one can hypothesize that while all verbs have a SCAN of −1, POSSEs vary in part because LANs vary in ways that can affect the data available to children.

As noted above, passive constructions already suggest that SCANs do not reflect external arguments of verbs. And like weather reports, “nominalizations” suggest SCANs of −1; cp. Chomsky (1970), Marantz (1984), Borer (2005), Ramchand (2008). One can cut to the chase, expect an onsides kick, or give someone a break. So prima facie, ‘chase’, ‘kick’, and ‘break’ fetch monadic concepts. One can instead posit linguistic “SCAN-reducing” operations that create nouns from verbs, passives from actives, etc. But this is hardly costless. And to repeat an earlier point, if a representation can be used to a representation that has a different valence, we have to consider the possibility that lexicalization is itself a reformatting process of this sort.

Still, even if lexicalized concepts are not fetched for purposes of semantic composition, they may leave “marks” on the resulting verbs. In particular, if ‘put’ lexicalizes a triadic concept, the verb may record this fact. And if for whatever reason, the introduced concept PUT(_) is especially unrestrictive—one can put on a happy face, put up with a colleague, put someone on, etc.—this may have lead speakers to regularly use ‘put’ with locative adjuncts. Over a few generations, this may have led children to conclude that while ‘put’ fetches a monadic concept of events, it also imposes a constraint on complex expressions in which the verb appears: such expressions should represent something partly in terms of how its location changed. (Recall that phrases like ‘in an hour’ can only modify “telic” predicates that represent endpoints of events.) The actual facts are surely more complicated. But one can certainly imagine that details concerning lexicalized concepts lead to patterns in usage that lead children to encode certain lexical restrictions on sentential expressions.

Relatedly, let’s distinguish at least two ways in which a string of words might fail to be (the string determined by) a grammatical expression. The contrast between (98) and (99)

(98) *Chris been have might there
(99) Chris might have been there

is apparently due to the limited ways in which expressions of a human I-language can be generated; see Chomsky (1957). The string ‘been might have there’ is not the “linearization” of any expression of (an I-language that is a version of) English. No procedure for assembling such an expression—forming ‘have might there’ and combining with ‘been’, forming ‘been have’ and ‘might there’ and combining, etc.—is a human way of generating expressions.²⁶ By contrast, while (100) is deviant for adult speakers of English, it is as generable as (100a).

(100) *Chris goed to the store
(100a) Chris went to the store.

²⁵ A lexical item that always fetches a singular concept has a SCAN of +1. A lexical item that sometimes fetches a monadic concept and otherwise fetches a dyadic concept has a SCAN of {−1, −2}.

²⁶ Recall that expressions are, or at least determine, articulation-meaning pairs. Obviously, a grammar can generate (an expression that determines) a word-string that sounds like ‘might been have there’ but transliterates the English expression ‘is easy to please’. But that is irrelevant.
The deviance of (90) lies with “irregularity” of a particular verb. Competent speakers know that the past tense of ‘go’ is ‘went’. In this sense, ‘goed to the store’ is generable but “filtered out.”

Recall that (101) has the meaning of (101a) and not (101b), the deviance of (101) differs from that of (98).

(101) *The child seems sleeping
(101a) The child seems to be sleeping
(101b) #The child seems sleepy

A competent speaker of English can somehow pair the articulation of (101) with the meaning of (101a). The deviance of (102) can make it hard to hear the meaning indicated with (102a).

(102) *Which linguist did the philosopher think that went home?
(102a) Which linguist did the philosopher think went home?

But one might think that (102) is a generable violation of a filter that precludes (pronunciation of) the complementizer ‘that’ when it immediately precedes a trace.27 I don’t want to take a hard stand on any particular case. My point is just that we can retain the traditional idea that a full characterization of (un)grammaticality will require appeal to some idiosyncratic constraints, at least for cases like (100), in addition to an intrinsically constrained generative procedure.

Ungenerability may be a “deeper” sort of ungrammaticality. But just as we’ve grown used to the idea that there are many sources of unacceptability, and often more than one way for a grammar to generate (an expression that determines) a certain string, we can and should get used to the idea that a grammar may fail to license a string that is determined by a generable expression.

In thinking about these issues, it is crucial that one not forget the reasons for adopting an I-language perspective on human languages. It is easy to forget that an I-language, consisting of a lexicon and a grammar, generates structured expressions that determine strings (relative to certain assumptions about how these expressions are linearized by relevant articulatory systems). It is easy to slide into thinking of languages as directly generating strings that either do or don’t meet “well-formedness” conditions of the sort specified when we specify familiar invented languages in the usual way. From this perspective (97) and (103) and are not expressions, and hence not meaningful expressions. But another option is to say (97) and (103) are generable but defective expressions: ‘send’ and ‘put’ are instructions to access concepts like SEND(_) and PUT(_): but ‘send’ also calls for specification of a patient; and ‘put’ calls for specification of a location. Likewise, one can say that (91) and (92) generable but defective.

(97) *Chris put the book.
(98) *Chris sent.

As Levin and Rappaport (1995, 2005) show, merely adopting a version of this second option would miss generalizations that we want to capture. But these may be generalizations about human concepts and how children tend to lexicalize them; and these facts can be described in terms of the variable POSSEs that children associate with lexical items, instead of assuming that the operations of semantic composition permit significant variation in LANs.

4. Number Neutrality
I have focused on adicity as an important dimension in which lexicalized concepts can differ from—and exhibit more typological variation than—introduced concepts that can be joined in simple ways. But I suspect that lexicalized concepts exhibit other typological distinctions that lexically introduced concepts do not, making the latter more systematically combinable than the

27 Note to Lasnik and ellipsis repair of island violations.
former. In particular, the concepts that children lexicalize may include both “mass concepts” and “count concepts;” where many of the latter are “singular,” in the sense of being applicable to a single thing, while others are “essentially plural.” But in my view, introduced analogs of these concepts are “number neutral,” and this neutrality—between mass and count, singular and plural—is related to the limitations on how human languages employ second-order quantification. Here, I will be applying and extending Boolos’ (1998) treatment of second-order quantification in an invented monadic calculus as plural quantification. The details will carry over into chapter six, which shows how phrases like ‘most of the cows that the dogs chased’ can be analyzed as instructions for how to build (second-order/plural) monadic concepts. I want to end this chapter by laying the ground for such analyses, while also showing that one need not focus on adicity to suspect that lexicalization is a process of using typologically diverse concepts to introduce less diverse concepts that can be systematically M-joined or Θ-joined. The idea will be that lexically introduced concepts are the atomic expressions of a mental language, SMPL, that is a second-order extension of a fragment of PL, the Tarskian language discussed in chapter two. And since SMPL only generates complex monadic concepts (a.k.a. mental predicates), it is far less powerful than CPL, Church’s lambda-extension of PL. But it is powerful enough to make it plausible that human linguistic meanings are instructions for how to build SMPL concepts.

4.1 Typology and Flexibility Again

I assume that many lexicalizable concepts are “mass concepts” in a familiar and much discussed sense. To think about some wood as wood is to think about it as some stuff of a certain sort, and thereby think about it in a way that does not involve thinking about it as (or even as constituted by things that are) countable. In thinking about anything as wood, we think about it in a way that lets us wonder how much of it there is in a given situation. If only for simplicity, I will also assume that in thinking about wood as such, we use a monadic/predicative concept as opposed to a relational concept. One can think that all wood is good, and recognize the implication that all brown wood is good; cp. §2.2 of chapter one. So I will speak of the mass-concept wood().

Other lexicalizable concepts are presumably “count concepts.” To think about a tree or some trees as such is to think about one or more things in a certain way—e.g., as a tree or as some trees—that lets us wonder how many of them there are in a given situation. At least many of these concepts are singular, or individualizing, in the following sense: if such a concept applies to anything, it applies to each one of the one or more things it applies to. The concept tree() applies to many things, but only in the distributive sense of applying to each of many trees. Each tree may comprise many things. But if tree() applies to any such things, taken together, that is because they constitute a tree; and while it may turn out that some trees are parts of other trees, tree() applies to some tree-parts only if each of them is itself a tree. By contrast, the complex concept three-trees() seems to be “essentially plural” in the sense of being potentially applicable to some things (taken together) but not to any single thing. Some atomic concepts may also be essentially plural; indeed, three() may be a such a concept.

A certain kind of theorist will say that three() is really a singular concept that applies distributively to certain “plural entities”—e.g., sets that have three elements. Likewise, one can hypothesize that trees() and three-trees() and even wood() are singular concepts that

---

28 Our concept of good(ness) may underlying relational; cp. §4.2 of chapter three. But our concept of wood seems to be classificatory. Though talk of a single concept, wood(), is idealization; see §1.2 of chapter one. In particular, a person who knows some science may have more than one concept of wood (water, fire, etc.). But I assume that we have a concept of wood that makes it possible to entertain the (false but coherent) thought that wood is basic stuff, not composed of anything else, and even that there is no smallest portion of wood. ADD REFERENCES.
apply to plural entities. I don’t deny that we can form concepts like \textsc{set-of-trees}() and \textsc{set-with-three-elements}() and \textsc{set-of-wood-particles}(). For reasons discussed below, I think we also form essentially plural concepts that are \emph{are not} singular concepts that apply to plural entities; and I think that many expressions, including ‘three trees’, are instructions for how to build such concepts. But for the moment, it doesn’t matter whether we distinguish singular concepts from plural concepts (and distinguish all of these from mass concepts), as opposed to drawing distinctions among singular concepts. The more important point is that just as lexicalized concepts seem to differ typologically in ways that we can describe with talk of adicity, they seem to differ along other dimensions, including mass/count and singular/plural; where these distinctions intended to describe kinds of \emph{concepts}, or \emph{ways of thinking about} things, not kinds of things that might be thought about in a uniform way; see §1.1 of chapter one.

I am not assuming that the world contains stuff that is not constituted by countable things. I am also not assuming that the world doesn’t contain any such stuff. (I am focused on meaning, not ontology.) But I do assume that humans and other animals often think about stuff-or-things \emph{as} countable things—or at least as things whose numerosity can be estimated—and that we often think about stuff-or-things \emph{without} thinking about it-or-them as countable. When we think about stuff-or-things \emph{as} countable, we sometimes do so via concepts that are essentially plural, and we often use concepts that are singular/individualizing/distributive.

However this familiar typology is described, the conceptual distinctions seem pretty robust. So as many authors have discussed, it is striking that linguistic expressions seem to be rather flexible in this respect. Examples like (104-106) abound.

(104) The red rock that is scattered in the yard was inexpensive.
(105) It came from a single red rock.
(106) There are more red rocks over there.

Similarly, a small portion of fish may have come from a fish that was among three fish caught this morning. One can say that ‘rock’ appears as a “mass noun” in (104) and as a pluralizable “count noun” in (105) and (106). But talk of mass/count/plural nouns may layers of grammatical structure, as opposed to typologically distinct kinds of lexical items; see, e.g., Gillon (1987), Borer (2005). I’ll return to the idea that just as ‘rocks’ is complex expression, not a lexical item, the count noun ‘rock’ is a kind of noun \emph{phrase} that combines a lexical \emph{root}, with a functional element. Unsurprisingly, I like this idea; the functional element can be viewed as an instruction to conjoin a concept like \textsc{one-or-more}() with a neutral concept, \textsc{\forall rock}(), which can be introduced and then fetched via the lexical item ‘rock’. But however we encode the distinction between the uses of ‘rock’ in (104) and (105), it seems clear that ‘rocks’ in (96) has the lexical item as a constituent. So let’s focus on rocks before turning to rock.

If only for simplicity of exposition, let’s assume that in acquiring ‘rock’, a child lexicalizes a concept \textsc{one-rock}() with which the child can think about a single rock as such. This leaves room for various hypotheses about (i) how the lexicalized concept is related to the concept(s) that can be fetched via the lexical item, and (ii) the role of the plurality marker. But given a Frege-Church conception of composition, an obvious first thought is that ‘rock’ fetches \(\lambda x. T \equiv \textsc{one-rock}(x)\)—a concept of type \(<e, t>\)—and the plurality marker fetches a concept of type \(<<e, t>, <e, t>>\) like \(\lambda x. \lambda p. T \equiv \textsc{plurality}(p) \& \forall x. \text{elementof}(x, p)[x x = T]\); where \textsc{plurality}() is a concept of plural entities, and the lambda concepts may have to be introduced, rather than simply labeled by infants who already function-denoters independent of pressure from lexicalization. Given the operation of function application, as discussed in chapter two, ‘rocks’ then be viewed as an instruction for a how to build a complex concept of plural entities—
perhaps collections or lattices—whose basic elements are countable rocks. This is an overtly mentalistic version of a familiar kind of account that is often presented in truth-theoretic terms; see Link (1983), Schwartzschild (1996), Chierchia (1998).

I don’t deny that infants have concepts like PLURALITY( ), which can be used to form concepts of pluralities of rocks. But it doesn’t follow that infants acquire a corresponding concept of \(<<e, t>, <e, t>>\) and use the plurality marker to fetch it. Another possibility is that ‘rocks’ is an instruction for how to build a concept like ROCK( )^PLURAL( ); where PLURAL( ) is an essentially plural concept that applies (nondistributively) to some things iff there are two or more of them, and ROCK( ) is a number-neutral concept that can apply to a single rock or to some rocks (however numerous). Put another way, ROCK( ) can be introduced so that it applies to one or more things iff ONE-ROCK( ) applies to each of them; cp. the number-neutral concept EACH-A-ROCK( ). I spell out this idea in section one of chapter six, drawing on Boolos (1998). But if it helps, think of ROCK( ) as akin to the concept ROCK-OR-ROCK-COLLECTION( ), which applies to entity e iff e is either a rock or a collection of rocks. Similarly, ROCK( )^PLURAL( ) is akin to ROCK-OR-ROCK-COLLECTION( )^PLURALITY( ), which applies to e iff e is a collection of rocks. However, kin differ. The concept ROCK( )^PLURAL( ) applies to rocks, not to collections.

I think we use ‘rocks’ to form concepts of rocks, not concepts of collections. Like Boolos, I don’t think (107) implies that there is a collection of the rocks in question,

(107) Theo saw the rocks.

much less that Theo saw any collection. There may be a collection of the relevant rocks, which may have been widely scattered, with each seen only by chance. There may also be a collection of those rocks and Theo, a collection of them and the event(s) of Theo seeing the rocks, etc. But I don’t think that (107)—or any thought built in accord with this instruction—implies that there is a collection of anything; though (107) does imply that there is at least one rock that Theo saw. Similarly, (108) can be used to make an assertion that implies the existence of various things,

(108) The red rocks that are scattered in the yard fell from the sky.

Yesterday, they rained down and formed a circle around a cactus.

but not a collection of rocks that rained down, formed a circle, and is scattered. And (109)

(109) Some logician iteratively specified the Zermelo-Frankl sets.

does not imply there was an iterative specification of a collection whose elements are the ZF sets.

One can imagine minds that use concepts like PLURALITY( ) to introduce higher-order concepts like \(\lambda X.\lambda p.T = \text{PLURALITY}(p) \& \forall x:\text{ELEMENTOF}(x, p)[xx = T]\), or perhaps use concepts like ONE-ROCK( ) to introduce concepts that apply to collections. But ascending the Fregean hierarchy types, or introducing concepts of abstractions, may be more of a logician’s move than a natural strategy for human children. And there is an alternative: given concepts that cannot be coherently M-joined or \(\Theta\)-joined, try to introduce concepts that can be so combined. The concept ONE-ROCK( ) applies to each rock, but PLURAL( ) does not. Simply M-joining these concepts, as if PLURAL( ) were like RED( ), would be incoherent. (I return to the mass-reading of ‘red rock’.) But if children have singular and plural concepts, independent of the I-languages they acquire, perhaps lexicalization is a process in which they acquire neutral concepts like ROCK( ). This concept can be coherently M-joined with PLURAL( ) to form a concept of rocks, not collections.

By hypothesis, a nonmonadic concept cannot be M-joined with anything. But the monadic concepts available to children, prior to lexicalization, may also exhibit typological distinctions that preclude systematic M-junction. There may be various ways of introducing more neutral concepts. Though one obvious (and perhaps biologically workable) strategy is to introduce concepts that are effectively disjunctive, taking advantage of the fact that predicate
reduction is a natural form of inference. If a child can use $\text{ONE-ROCK}(\_)$ to introduce $\#\text{ROCK}(\_)$, given some ancillary resources, then it might be easy to form concepts like $\#\text{ROCK}(\_) \wedge \text{PLURAL}(\_)$ and $\#\text{ROCK}(\_) \wedge \text{ONE}(\_)$; where is $\text{ONE}(\_)$ effectively equivalent to $\neg \text{PLURAL}(\_)$, and so $\#\text{ROCK}(\_) \wedge \text{ONE}(\_)$ is effectively equivalent to $\text{ONE-ROCK}(\_)$. Correlatively, the resources required to introduce $\#\text{ROCK}(\_)$ are relatively modest.

Suppose that a child acquiring an I-language has a concept $\text{INCLUDE}(\_, \_)$ that can be used to think that one or more independently specified things include an independently specified thing. Thinking that Bessie is a cow is not very different than thinking that the one or more cows include Bessie, or that Bessie is among the one or more cows. Indeed, $\text{INCLUDE}(\_, \_)$ can be viewed as the inverse of $\text{ONE-OF}(\_, \_)$; each cow is one of the cows, each of which is a cow. If it helps, think of $\text{ONE-OF}(\_, \_)$ as akin to $\text{ELEMENT-OF}(\_, \_)$. But my suggestion is that when children are introducing fetchable concepts, in the course lexicalization, they can use concepts like $\forall: \text{INCLUDE}(\_, \_)[\text{ONE-ROCK}(\_)]$; where this complex monadic concept applies to one or more things iff each of them is a rock. For some purposes, this might count as a variant of the idea that children can introduce the concept $\lambda X.\lambda P.\text{T} \equiv \text{PLURALITY}(P) \& \forall X: \text{ELEMENT-OF}(X, P)[Xx = T]$, fetch it via the plural marker, and saturate it with the concept $\lambda X.\text{T} \equiv \text{ONE-ROCK}(X)$. But here, the issues concern the range of human linguistic types and the nature of composition operations.

Recall that $\Theta$-junction permits construction of $\exists[\text{UNDER}(\_, \_) \wedge \text{ONE-ROCK}(\_)]$, which applies to things that are under a rock. This monadic concept is akin to the Tarskian open sentence $\exists X[\text{UNDER}(X', X) \& \text{ONE-ROCK}(X)]$. But if the “slots” in $\exists[\text{UNDER}(\_, \_) \wedge \text{ONE-ROCK}(\_)]$ are number neutral, and $\exists[\text{FORM}(\_, \_) \wedge \text{ONE-CIRCLE}(\_)]$ can apply to some things that together form a circle, then Tarskian sentences are both too sophisticated and too limited as models for these $\Theta$-junctions: too sophisticated because they impose no limit on the number of open variables in a conjunct; too limited, because each variable has exactly one value relative to each assignment of values to variables. And I don’t want to say that $\exists X[\text{FORM}(X', X) \& \text{ONE-CIRCLE}(X)]$ is a concept of sets. Likewise, I am positing concepts like $\forall: \text{INCLUDE}(\_, \_)[\text{ONE-ROCK}(\_)]$ instead of positing concepts like $\forall X: \text{INCLUDE}(X', X)[\text{ONE-ROCK}(X)]$ and saying that they true of certain sets that are extensions of functions of type $<e, t>$.

### 4.2 Two Ways of Interpreting Capitalized Variables

Luckily, there is a better model. The trick is to allow for quantification into predicate positions without thereby quantifying over anything new.

One can invent a Frege-Church language that generates (110-112),

\[\begin{align*}
(110) \ Pa \\
(111) \ \exists x[Px \ & (x = a)] \\
(112) \ \exists x[Xa \ & \ \forall x(Xx = Px)]
\end{align*}\]

with each of these sentences being equivalent to the others. If ‘a’ denotes an entity, and ‘P’ denotes a function from entities to truth values, then (110-112) are mutual paraphrases. One can also invent a variant of the Tarskian language PL that generates the same strings, but still has no denoters, letting (112) be shorthand for (113). Then (114) can be a theorem while (115) is not.

\[\begin{align*}
(113) \ \exists s:\text{Set}(s)[a \in s \ & \ \forall x(x \in s \equiv Px)] \\
(114) \ Pa \equiv \exists x[Px \ & (x = a)] \\
(115) \ Pa \equiv \exists x[Xa \ & \ \forall x(Xx = Px)]
\end{align*}\]

Even if (115) is true, its right side has implications like (115a) and (115b);

\[\begin{align*}
(115a) \ \exists s:\text{Set}(s)[a \in s] \\
(115b) \ \exists s:\text{Set}(s)[\forall x(x \in s \equiv Px)]
\end{align*}\]

31
and one can stipulate that these are not implications of (110) or (111) in the variant of PL.

In the second kind of language, the “second-order” quantification in (102) and (105) is disguised first-order quantification over sets. Quine (196x) and others have asserted that eschewing second-order quantification in favor of explicit first-order quantification over sets—or perhaps merological sums, collections, lattices, or other “potentially plural” entities that can somehow “include” more basic entities—is virtuous in some way that matters for science. I am not persuaded. But whatever one thinks about the languages that mathematicians and physicists (along with linguists and philosophers) should use, for scientific purposes, there is no reason for thinking that human children have been acquiring such languages for millennia.

Humans, perhaps along with some other animals, may enjoy one or more languages of thought that permit quantification into “predicate positions.” This is not to say that children can formulate mental analogs of (112), which already includes an ampersand along with a first-order quantifier that can appear in sentences like (116).

\[(116) \forall x (Rxx' \& Sx'x'x''')\]

But children might be able to formulate mental predicates that are more like (117),

\[(117) \exists [\text{AGENT}(, ,)^\text{DOGS}(\_) \text{CHASE}(\_)^\text{PAST}(\_) \exists [\text{PATIENT}(, ,)^\text{CATS}(\_)]\]

which is a concept of past chases whose agents are some dogs, and whose patients are some cats. If the unsaturated slots are like Tarski-style first-order variables, then any assignment of values to the slots would assign exactly one entity to each slot, perhaps allowing sets as special entities. If the unsaturated slots are like Church-style second-order variables, then any assignment of values to the slots would assign exactly one function to each slot. For many purposes, including mine, that is an unimportant contrast. But a third interpretation for (112)

\[(112) \exists X [Xa \& \forall x (Xx \equiv P)]\]

provides a helpful model for the slots in (117).

Boolos offers an interpretation according to which second-order variables are formally distinctive while having values of the same sort as first-order variables. Instead of treating a capitalized variable like ‘X’ as a disguised first-order variable, ranging over sets or other abstracta that bear a special relation to values of ‘x’, Boolos treats each capitalized variable as a number-neutral variable that can have one or more values relative to each assignment of values to variables. Each lower case variable is still construed as a singular variable that always has exactly one value relative to each assignment of values to variables.

On this interpretation, ‘Xx’ indicates that the value of ‘x’ is one of the one or more values of ‘X’. Or put the other way, the one or more values of ‘X’ include the one value of ‘x’. For any two or more things (the Xs), each thing (x) bears the “one of” relation to them iff it is (identical with) one of them. Each thing also bears this relation to itself, but to no other single thing: if the one or more Xs are just one, then trivially, that lone X is still one of the one or more Xs; it is the X. On this view, to say that Aristotle is one of the one or more philosophers is not to say Aristotle is an element of a certain set, or that he is mapped to a truth value by a certain function. The claim is simply that there are one or more philosophers, and Aristotle is one of them. Likewise, ‘\(\forall x (Xx \equiv P)\)’ indicates that each domain entity is one of the Xs iff it is one of the Ps. If these values of ‘P’ are the philosophers, and ‘a’ indicates Aristotle, then (102) says that one or more things are such that Aristotle is one of them, and each thing is one of them iff it is a philosopher. Which is just to say, elaborately, that Aristotle is a philosopher.

In English, one says ‘one or more things are’, not ‘one or more things is’. But nonsingular agreement does not imply more than one thing. Consider (118) and (119).
(118) There are one or two philosophers here.
(119) There are zero/between zero and two/1.0/no philosophers here.
And while (120) implies that the room contains more than one philosopher, (121) does not.
(120) There are philosophers in the room, and he is one.
(121) There are one or more philosophers in the room, and he is one.
Grammatical agreement cannot keep us from using ‘one or more’ number neutrally. And a
tradition of set-theoretic construal cannot force us to interpret ‘X’ as having exactly one value
relative to each assignment of values to variables. We can interpret ‘X’ as a variable that can
have many values, none of which need to be sets.
This is to deny that an assignment must be characterizable as a function that maps each
variable to exactly one entity. And this is to make the notion of an assignment more general than
Tarski’s original notion. But there is nothing mysterious about the more general notion. On the
contrary, the apparatus usually invoked to make the set-theoretic construal explicit can be used to
make the Boolos-construal equally explicit. For illustration, assume a domain of exactly 4 things.
This yields 15 potential assignments, which can be depicted in either of the ways shown below.

<table>
<thead>
<tr>
<th>abcd</th>
<th>0001</th>
<th>0010</th>
<th>0011</th>
</tr>
</thead>
<tbody>
<tr>
<td>abc</td>
<td>abd</td>
<td>acd</td>
<td>bcd</td>
</tr>
<tr>
<td>ab</td>
<td>ac</td>
<td>ad</td>
<td>bd</td>
</tr>
<tr>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
</tr>
</tbody>
</table>

The representation on the left can be used to illustrate the set-theoretic construal: the bottom
nodes of the “lattice” indicate “basic” elements of an extended domain; every other node
indicates a non-primitive entity that can be the value of a second-order variable. Then ‘abd’
signifies the set \{a, b, d\}, or the “mereological sum” a⊕b⊕d. But instead of introducing an
extended domain, one might take ‘abd’ to indicate a plural assignment of three values to a
second-order variable. Recoding in base two—a = 1, b = 10, c = 100, d = 1000—makes this
explicit. For while ‘1011’ can be viewed as a label for a thing that has three elements, the binary
code can also be read as providing answers to yes/no questions. For each entity—d, c, b, and a—we can ask whether it is one of the potentially many values assigned to the variable, with ‘1011’
providing the following answers: yes, no, yes, and yes.
Given this apparent equivalence, one might wonder if the plural interpretation of ‘X’
differs at all from the set-theoretic interpretation. But consider the infamous formula (122),
(122) \(\exists X \forall x [Xx \equiv (x \notin x)]\)
with ‘x \notin x’ meaning that the value of ‘x’ is not selfelemental. And let’s stipulate a domain that
is exhausted by two things—say, Aristotle and Frege—neither of which is an element of itself.\footnote{Chomsky can be the only linguist, and so one of the one or more linguists, without being selfelemental. Insisting that each thing is a “part of” itself is, I think, to interpret ‘part’ (misleadingly) as ‘one’. A related point: competent speakers can understand and assent to the sentence ‘There are many sets, each of which is a set, but there is no set whose elements are all and only the sets’. This suggests that speakers need not assume that each noun corresponds to the set of things satisfied by that noun. Likewise, speakers can understand and assent to ‘No set contains all and only the sets that do not contain themselves’, just as they can understand and assent to ‘No barber shaves all and only the barbers who do not shave themselves’.}
Given such a domain, (122) is true on the Boolos interpretation: there are some things such that
for each thing, it is one of them iff it is not selfelemental. But since nothing in the stipulated
domain has each thing as an element, (122) is false on the more standard interpretation. Given a
different domain that includes \{Aristotle, Frege\} along with its two elements, (122) can be true
on both interpretations. But the interpretations still differ. For there are many domains relative to
which (122) is true on the Boolos construal but false on the set-theoretic construal. Consider the
Zermelo-Frankl (ZF) sets. Relative to this familiar domain: there are some things—namely, the
ZF sets—such that each thing (i.e., each ZF set) is one of them iff it is not selfelemental; yet
there is no set whose elements include all and only the nonselfelemental things. So even if other
domains include a “collection” of all and only the ZF sets, at least one perfectly standard domain
makes (122) true on one interpretation but false on the other. So the interpretations differ.

Correlatively, (123) is trivial on the Boolos interpretation.

\[(123) \exists x(Fx) \equiv \exists x\forall x(x \equiv Fx)\]

Given an F, there are one or more Fs—i.e., one or more things are such that each thing is one of
them iff it is an F—and given one or more Fs, there is an F. This does not imply that the Fs form
a set, although they may well. Likewise, (124) is trivial.

\[(124) \forall x \sim \exists x(Xx & \sim Xx)\]

For any (choice of) one or more things, nothing is both one of them and not one of them. Or put
another way, replacing ‘\(\forall x\)’ with ‘\(\sim \exists x\)’ and eliminating the double negation:

\[(124a) \sim \exists x\exists x(Xx & \sim Xx)\]

it’s false that one or more things are such that something is both one of them and not one of them.
These formulations of monadic second-order noncontradiction are not claims about sets. If it
helps, imagine replacing the singular variable ‘x’ with ‘\(\theta\)’, and the number-neutral ‘X’ with ‘\(\Theta\)

\[(124b) \sim \exists \Theta \exists \Theta[OneOf(\Theta, \Theta) \equiv \sim OneOf(\Theta, \Theta)]\]

By contrast, (125/125a) is true, letting ‘s’ be a special variable ranging over sets.

\[(125) \forall s \sim \exists x[(x \in s) & \sim (x \in s)]\]

\[(125a) \sim \exists s \exists x[(x \in s) & \sim (x \in s)]\]

But this seems more like a special case of noncontradiction: for any set, no thing both is and isn’t
an element of it. In this sense, (125) seems like (126).

\[(126) \forall x \sim \exists x[In(x, y) & \sim In(x, y)]\]

As Boolos notes, his interpretation does not extend to second-order relations that resist
characterization as predicates of n-tuples. But instances of (127) can be glossed as in (127a).

\[(127) \exists R...Rxy…\]

\[(127a) one\ one\ ordered\ pairs,\ the\ Rs,\ are\ such\ that…they…\]

And even with just a (very) finite number of dyadic relations, including some thematic relations,
this yields a lot of expressive power—perhaps enough, or even more than enough, to adequately
model the concepts that can be constructed by executing the assembly instructions that human
linguistic expressions connect with pronunciations.

One can hypothesize that each human language is a fragment of a language like PL or
CPL. Though if one suspects that sentences of a human language do not themselves have truth
conditions, one might say that meanings are instructions for how to build expressions of a mental
language that is a fragment of like PL or CPL. My suggestion is that meanings are instructions
for how to build expressions of a mental language—call it SMPL—that is more-second order
than PL, but massively monadic. While SMPL allows for finitely many atomic dyadic concepts,
every complex concept of SMPL is monadic; although many atomic concepts of SMPL may be
introduced by using available polyadic concepts. And while SMPL is like CPL in allowing for
quantification into predicate positions, as in the \(\Theta\)-junction \(\exists[AGENT(\_, \_)^{\sim DOGS(\_)},\]
the concepts generated by SMPL are predicative, not denoters; although some atomic concepts of SMPL may be introduced by using available denoting concepts.

Moreover, the monadic concepts generated by SMPL are not instances of the Fregean type \(<e, t>\). Concepts of the SMPL type \(<M>\) can be described from a Fregean perspective; but they do not themselves relate entities to truth values. Likewise, \(\exists[\text{AGENT}(\_ , \_ )^{\text{DOGS}(\_ )}]\) is not an instance of the Fregean type \(<e, t>\); \(\text{AGENT}(\_ , \_ )^{\text{DOGS}(\_ )}\) is not an instance of the Fregean type \(<e, \langle e, t \rangle >\); and \(\exists\) is not an instance of the Fregean type \(\langle e, \langle e, t \rangle >, \langle e, t \rangle >\). There is an operation of \(\Theta\)-junction that permits combination of atomic dyadic concept with a monadic concept; and since the combined concepts may be plural, the kind of existential closure involved is better viewed as number neutral, a la Boolos, than as a Church-style second-order quantifier (or a Tarski-style first-order quantifier) whose variables range over sets of some kind.

If the slots in thematic concepts are characterized number-neutrally, we can say that \(\text{AGENT}(\_ , \_ )\) applies to some entities and some events iff the latter are (all and only) the agents of the former; cp. the discussion of thematic “exhaustion” in §3.2 of chapter three. Following Schein (1993, 2002, forthcoming), I assume that ‘branded fifty cows with twenty irons’—which can combine with ‘Five ranchers’—can be understood distributively, implying that each brander branded fifty cows, or collectively. The collective reading, which implies only fifty brandings, implies nothing about how the cows and irons were distributed across those who got the work done. And this neutrality is captured if each thematic role has its own number-neutral conjunct, like \(\exists[\text{AGENT}(\_ , \_ )^{\langle\text{FIVE}(\_ )\text{RANCHERS}(\_ )\rangle}]\), according to which five ranchers did the work.

Likewise (128) exhibits a “semi-distributive” reading:

(128) Three trainers taught five dogs four tricks yesterday

which three trainers together taught four tricks to each of five dogs, for a total of twenty processes of teaching a dog a trick—perhaps with more than one trainer involved in a given trick-teaching, but with no implication (or prohibition) of cooperation. As Schein (1993, 2002) argues, such facts tell against analyses in terms of polyadic concepts like \(\text{TEACH}(e, X', X', X')\) that are satisfied by \(n\)-tuples of collections of events/trainers/dogs/tricks. Though even setting such examples aside, relative clauses like ‘who branded fifty cows with twenty irons yesterday’ apparently correspond to concepts that can be satisfied by one individual who did a lot of work, or by several individuals who together did the requisite branding (with or without cooperation).

Even if ‘dogs chase cats’ can be analyzed as instruction for how build (117),

(117) \(\exists[\text{AGENT}(\_ , \_ )^{\text{DOGS}(\_ )}]^{\text{CHASE}(\_ )^{\text{PAST}(\_ )}}\exists[\text{PATIENT}(\_ , \_ )^{\text{CATS}(\_ )}]\)

that is still a long way from making it plausible that meanings are more generally instructions for how to build SMPL concepts. But idea is to figure out what must be added to \(M\)-junction and \(\Theta\)-junction, given a suitable stock of potentially plural atomic monadic and dyadic concepts, to generate the meanings of human I-language expressions—instead of starting with a far more powerful model like CPL, and asking what must be subtracted to not overgenerate meanings.

### 4.3 Mass without Church

If children can form concepts like \#\text{ROCK}(\_ ), which can apply to one thing or to many things, that may help with (105) and (106), but it doesn’t yet accommodate the use of ‘rock’ in (104).

(104) The red rock that is scattered in the yard was inexpensive.

(105) It came from a single red rock.

(106) There are more red rocks over there.

Perhaps the lexical item ‘rock’ is used to fetch either \#\text{ROCK}(\_ ) or a concept of certain sets whose elements are parts of rocks. But if so, then a certain simplicity of SMPL is lost. If \text{ONE-ROCK}(\_ ) can be used to introduce \text{SET-OF-ROCK-PARTS}(\_ ), which can then be fetched and
conjoined with a concept that can be fetched via ‘red’, then it would seem that SET-OF-ROCKS(\_\_) can be introduced, fetched, and conjoined with a concept that can be fetched via ‘s’. In which case, SMPL starts looking like a fragment of CPL. Of course, the concepts we build by executing meanings are what they are. But perhaps Boolos’s strategy can be extended to allow for predicates and quantifiers that are mass/count neutral.

Let’s return the idea that the lexical item ‘rock’ is a constituent of the count noun ‘rock’, which is a constituent of the plural noun ‘rocks’. If the corresponding concept ROCK(\_) is the complex concept #ROCK(\_)\^PLURAL(\_)—with #ROCK(\_) as an introduced concept that is neutral enough to be coherently conjoined with either PLURAL(\_) or ONE(\_)—perhaps #ROCK(\_) is the complex concept √ROCK(\_)\^ONE-OR-MORE(\_); where √ROCK(\_) is an introduced concept that can but need not be conjoined/restricted with a concept that is used to think of stuff-or-things as one or more things. But let me stress at once: the idea is not to analyze ONE-ROCK(\_) in terms of a more basic “root concept” that applies to a certain kind of stuff as opposed to countable things.

I am not suggesting that all count concepts are complex concepts that have concepts of stuff as constituents. In my view, thinking of a cow as such is not a matter of thinking of something as a certain hunk of beef with some interesting properties not exhibited by beef on the plate. On the contrary, thinking of some beef as such is more likely a matter of thinking about it as meat from one or more cows. Likewise, thinking about a pound of rock, fish, or chicken as such seems to be a matter of thinking about it as a pound of stuff from rocks, fishes, or chickens. We need to leave ample room for the idea that √ROCK(\_) is introduced in terms of ONE-ROCK(\_), √FISH(\_) is introduced in terms of ONE-FISH(\_), etc. In my view, a singular count noun can have a lexical root that is used to fetch a number-neutral concept that is introduced via some concept that is equivalent to the concept constructed via the count noun. If ‘a fish’ is the structured expression [[fish+CT]–PL]—where ‘CT’ and ‘PL’ are grammatical features of count and plurality—then the atomic expression ‘fish’ can use a root concept √FISH(\_) that can be conjoined with ONE(\_)—or with ~PLURAL(\_) and ONE-OR-MORE(\_)—to form a concept that is equivalent to an atomic concept, ONE-FISH(\_), in terms of which √FISH(\_) was introduced.30

Talk of a “universal grinder” (see Pelletier 198x) can find its place here, without the usual Fregean apparatus and appeals to shiftable types. I think √FISH(\_) can be introduced as a concept that applies to some stuff-or-thing(s) iff ONE-FISH(\_) applies to all of them or all of it is the result of “grinding” things that ONE-FISH(\_) applies to. I’ll try to make this idea a bit more precise below. But the details are less important than recognizing that there are cases and cases. To take the usual contrasting example, thinking of some water as such is presumably not a matter of thinking about it as stuff from a water. Our concept of a water is presumably introduced via some mass concept of water and some “unitizing” concept.

Given a mass concept WATER(\_) that we can use to think about water as such, we can form concepts like SUITABLE-PARTITION-OF-WATER(\_). I have nothing interesting to say about how to form such concepts; cp. the discussion of EAT(\_) and DINE(\_) in §3.x above. But given that we can form the concept of a water, I think we can form the neutral concept √WATER(\_), which applies to both whatever the mass concept WATER(\_) applies to and to waters, whatever they turn out to be. Though in most contexts, it may be clear that the “peunembral expansion” of

---

30 Chierchia (201x) describes the meanings of (at least many) mass nouns in terms of “uncountable countables,” using (un)countability as a psychological notion. I’m not sure how this fits with formulating meaning theories as truth theories. But I agree with Chierchia—along with many others, including Gillon (1987), Bale and Barner (2009)—that representing some stuff-or-things as countable is an aspect of representation in addition to merely representing the stuff-or-things as such.
WATER(_ ) is irrelevant, and that one is to ignore any waters that √WATER(_ ) applies to.

In recognizing that there are cases and cases, we should also be wary of misdiagnosing beliefs about the nature of the stuff-or-things we talk about as semantic implications. Most of us do not believe that there farms on which tofus are raised and slaughtered. But imagine a child whose parents are conspiracy theorists. They regale the child with stories of tofus running free before letting the child see any tofu, and then tell the child that eating tofu is as bad as eating chicken. Such a child might acquire an atomic concept of a tofu, ONE-TOFU(_ ), and use it to introduce the more permissive concept √TOFU(_ ), which can be restricted to form a complex concept that is equivalent to ONE-TOFU(_ ). But our concept ONE-TOFU(_ ) may be complex and contrived, somehow built up from √TOFU(_ ), which we introduce via some mass concept that we use to think about tofu without ever thinking of it as countable.

I don’t see any reason for thinking that human I-languages somehow encode the facts about tofu(s), chicken(s), fish, sheep, rock, paper, scissors, etc. It seems more likely that we use whatever concepts we start with to introduce more neutral concepts that we can then use in thinking that there are no tofus, and that the poor child has been misled about what tofu is—or that there are tofus, and that the general public has been misled about where tofu comes from. A child might acquire an atomic concept ONE-SKY(_ ), perhaps after hearing utterances of sentences like (128), and introduce √SKY(_ ) to accommodate sentences like (129).

(128) The sky is blue today. But skies here are never as blue as skies in Arizona.
(129) There is a lot of blue sky today. There was much less blue sky yesterday.

But given my conception of meaning, another child might acquire an atomic mass concept SKY(_ ) and introduce √SKY(_ ) to accommodate sentences like (128). Recall the different perspectives on the concepts NUMBER(_ ) discussed in §1.x of chapter two: we humans may naturally acquire an atomic monadic concept of numbers that we can use to think about the number/cardinality of a certain set; but ideal Fregean thinkers might view our concept NUMBER(_ ) as mental shorthand for a more complex higher-order concept.

Correlatively, we can find a place for talk of “coercion” in an account of which concepts children are apt to lexicalize, without supposing that these concepts are semantically privileged in a way that requires talk of Fregean types and type-shifting. Moreover, while I take lexical items to be atomic expressions of a human I-language, I do not assume that lexical items are the only meaning-pronunciation pairs that speakers store in some form of long-term memory. I think the meaning of the lexical item ‘chicken’ is the instruction fetch@’chicken’, whose execution accesses a neutral concept √CHICKEN(_ ) that applies to chickens, each chicken, and each portion of chicken. My suspicion is that for most speakers, this neutral concept is introduced by using an atomic count-concept, ONE-CHICKEN(_ ). Once introduced, √CHICKEN(_ ) can be combined with other concepts to form a complex concept that is analytically equivalent to ONE-CHICKEN(_ ). This provides a kind of explanation for how we know that chicken comes from chickens. But this does not imply that each time we hear ‘a chicken’, we go through the bother of constructing √CHICKEN(_ )∧ONE(_ ). It may be more efficient to maintain a “second address” for the complex expression [chicken+CT], so that ONE-CHICKEN(_ ) can be accessed in contexts where the more complex analytically complex concept is not needed. This is a kind of homophony, but not lexical homophony. And if it seems bizarre to even think of “grinding” dots to create a gallon of

31 As opposed to some mass-concept and a “unitizer.” But I once heard an anecdote from a mother reporting an awkward conversation, in which her daughter at first thought it was funny that the stuff they were eating for dinner “had the same name as” the chickens they had seen the day before at the petting zoo.
dot that can be spread all over the wall, then we may need to be “coerced” into forming $\sqrt{\text{DOT}(\_)}$ and using fetch this neutral concept to think about some stuff as a gallon of dot.

If we never encounter the expression [dot–CT], but often encounter [dot+CT], we might connect the pronunciation of ‘dot’ with an instruction of the form $\text{fetch@L}$, where $\text{ONE-DOT(\_)}$ is the only concept at L. But forcing a speaker to deal with [dot–CT] may trigger a process of creating a new entry for the lexical item itself, so that ‘dot’ comes to have two related addresses, corresponding to two related instructions: $\text{fetch@\textquoteleft dot\textquoteright}$, whose execution accesses $\sqrt{\text{DOT}(\_)}$; and $\text{fetch@\textquoteleft dot+CT\textquoteright}$, whose execution accesses $\sqrt{\text{DOT}(\_)\text{\^\_\text{ONE}(\_)}}$. As usual, we need to distinguish the elements and operations of an I-language from systems that supplement a basic I-language in ways that facilitate its various uses.

Still, one would like a model for posited concepts like $\sqrt{\text{ROCK}(\_)}$ and $\sqrt{\text{FISH}(\_)}$. Boolos showed how the capitalized variable in ‘$\forall y[\text{OneOf}(y, X) \equiv \text{Fish}(y)]$’ can be interpreted as having one or more values relative to a single assignment of values to variables. The invented formula is true of one or more things, the Xs, iff each thing, y, is such that it$_y$ is one of them$_X$ iff it$_y$ is a fish. To extend this singular/plural neutrality in a way that also yields mass/count neutrality, we need to allow for assignments of value—like some mud or some beef—to a variable, without insisting that each assignment assign one or more values to each variable. Alas, we ‘some fish’ is ambiguous. So lets use ‘a fish’ for the singular expression and use ‘sm fish’ for the mass/plural expressions. We can disambiguate the latter by speaking of sm fish that is not a fish, or sm fish that are not fish. The idea, then, is to relax the Boolos interpretation of variables so a single assignment of value-or-values to variables can assign to a single variable sm value that is not a value (e.g. sm fish that is not a fish) or one or more values (e.g. one or more fish).

I am not suggesting that this is a good idea for purposes of studying the relation between logic and arithmetic, which is intimately connected with notions of counting countable things. Correlatively, it may not be a good idea to impose a Fregean logic—rooted in notions of mapping countable things onto truth values—on mass concepts or number-neutral concepts. But in any case, my proposal concerns mental representations, not the stuff-or-things represented. Given a domain of countable things, Boolos showed how ‘OneOf(y, X)’ can be satisfied by an assignment $\mathcal{A}$ that assigns exactly one thing to ‘y’ and one or more things to ‘X’: the thing that $\mathcal{A}$ assigns to ‘y’ just has to be one of the thing(s) that $\mathcal{A}$ assigns to ‘X’. But I think we can generalize this idea and say that ‘SomeOf(y, X)’ is satisfied by assignment $\mathcal{A}$ iff whatever $\mathcal{A}$ assigns to ‘y’—be it a value or sm value that is not a value—is some of whatever $\mathcal{A}$ assigns to ‘X’; where $\mathcal{A}$ may assign to ‘X’ one or more values, or sm value that is not a value.

The formula ‘$\exists X \forall y[\text{SomeOf}(y, X) \equiv \text{Fish}(y)]$’ can be read as follows, using ‘be’ instead of ‘is or are’: there be some (stuff or thing or things), the X, such that there be some (stuff or thing), y, such that it$_y$ is some of it-or-them$_X$ iff it$_y$ be fish. As a special case of this “pirate logic” interpretation, consider an assignment $\mathcal{A}$ such that: $\mathcal{A}$ assigns a certain 2-pound fish, and nothing else, to ‘X’; and $\mathcal{A}$ assigns the co-located 2 pounds of fish, and nothing else, to ‘y’. Absent good

---

32 Likewise ‘$\exists X \forall Y[\text{OneOrMoreOf}(Y, X) \equiv \text{Fish}(Y)]$’ is true iff there are one or more things, the Xs, such that any one or more things that are plural, the Ys, are such that they$_Y$ are one or more of them$_X$ iff they$_Y$ are (i.e., iff each of them$_Y$ is a) fish. And ‘$\exists X \forall Y[\text{OneOrMoreOf}(Y, X) \equiv \text{Fish}(Y)]$’ is true iff there are one or more things, the Xs, such that any one or more things that are not plural, the Ys, are such that they$_Y$ are one or more of them$_X$ iff they$_Y$ are fish. So at least in principle, the lower case variables are dispensible.
reasons for distinguishing the fish that \( \mathcal{A} \) assigns to ‘X’ from the fish that \( \mathcal{A} \) assigns to ‘y’, I am inclined to identify the former fish with the latter fish, and diagnose this as a special kind of Hesperus/Phosphorus case. With regard to the domain of things-or-stuff, the (2 pounds of) fish can be the (2 pounds of) fish, even if you think of it both as a live animal that can gain some weight and as some food, of a certain quantity, that can cease to be a live animal. Likewise, I think (130) (130) The fish is on the table.
can be used to describe on which there are three (whole) fish, so long as we’re thinking about those fishes as dinner—or for some other use where counting is not important, and there might as well have sm fish that are not fish on the table. But if there are exactly three fish on the table, I am inclined to identify the fish that is on the table with the fish that are on the table.

That said, I am not committed to not distinguishing the fish from the fish; if there are good reasons for distinguishing the animal(s) from the corresponding matter, so be it. I tend to be suspicious of arguments that purport to extract ontological conclusions from modal premises—or premises concerning how we use ‘is’ as opposed to ‘are’—absent independent reasons for thinking that the premises reflect how things are, as opposed to how we think about them. But I’m no metaphysician. My goal is offer a plausible theory of meaning that can be incorporated into a plausible account of human nature and its place in nature.

Other things equal, I prefer theories of meaning that don’t have tendentious implications for what there is in nature, independent of human nature. But for those who want meaning theories to take the form of truth theories, it would seem especially important that ontological distinctions be motivated by considerations independent of linguistic meaning or conceptual content. In any case, the question of whether or not whatever \( \mathcal{A} \) assigns to ‘X’ differs from whatever \( \mathcal{A} \) assigns to ‘y’—like the question of whether or not the morning star differs from the evening star—doesn’t seem to be a question about human I-languages. So given that humans are apt to mischaracterize intentional distinctions as ontological distinctions, I think semanticists should be especially cautious when making assumptions about the stuff/things that can be talked about. In any case, my use of ‘SomeOf(y, X)’ does not presuppose that whatever \( \mathcal{A} \) assigns to ‘X’ differs from whatever \( \mathcal{A} \) assigns to ‘y’.33

4.4 Summary
While the details in this chapter have been many, the moral is simple. Lexical items are remarkably flexible in ways that tell against the idea that lexicalization is process in which linguistic expressions label concepts and inherit their adicities. Many different facts suggest that lexicalization is a process in which available concepts are used to introduce formally new concepts that can be combined via operations that take monadic or dyadic concepts as inputs.

33 Likewise, my use of ‘~Plural(Y)\OneOrMoreOf(Y, X)’ does not presuppose that that whatever \( \mathcal{A} \) assigns to ‘X’ differs from whatever \( \mathcal{A} \) assigns to ‘Y’. And we can, in principle, dispense with lower case variables (see note 33). Let ‘\( \exists X \forall Y [\SomeOf(Y, X) = \text{Fish}(Y)] \)’ be true iff there be some (stuff or thing or things), the X, such that any (i.e., any stuff or one or more things), the Y, be such that it-or-they\( Y \) be some of it-or-them\( X \) if it-or-they\( Y \) be fish. Restricting to countable things, ‘\( \exists X: \OneOrMore(X) [\forall Y: \OneOrMore(Y) [\SomeOf(Y, X) = \text{Fish}(Y)]] \)’ is true iff there be one or more, the Xs, such that any that be one or more, the Ys, be such that it-or-they\( Y \) be among it-or-them\( X \) iff it-or-they\( Y \) be fish. It might be said that since two things could have been assigned to ‘X’ but not to ‘y’, \( \mathcal{A}(\text{‘X’}) \) could have been two while \( \mathcal{A}(\text{‘y’}) \) was one, and so \( \mathcal{A}(\text{‘X’}) \) is distinct from \( \mathcal{A}(\text{‘y’}) \). But this is either (i) a crazy inference, or (ii) a way of introducing a notion of distinctness according to which “things” are individuated relative to a kind of variable. CITE JPL ARTICLE.
Chapter Six: Minimal Semantic Instructions

Most of the last chapter was devoted to describing the concepts that are fetched and assembled, on my view, when meanings are executed. In this chapter, I provide some compositional details regarding the instructions themselves. My plan is to sketch a familiar kind of “minimalist” account of how human linguistic expressions are generated, show how these expressions can be viewed as instructions for how to assemble concepts, and specify the instructions in enough detail to make it clear how they can be executed for some textbook constructions. At least for these constructions, I think my account preserves the main insights that have been obtained by developing truth-theoretic conceptions of linguistic meaning. My hope is that working through some elementary cases will also indicate how insights regarding other constructions—described in the vast literature that describes the meanings of human I-language expressions in truth-theoretic terms—can be redescribed in the “instructionist” terms described here.

In section one, I discuss untensed clauses like (1) and (2),

(1) a spy stab a colonel with a knife
(2) see Scarlet stab Mustard

in order to show how such clauses can be instructions for how to build monadic concepts of events or states that have participants, some which can be described as “external participants” (e.g., agents) or “internal participants” (e.g., patients). Section two focuses on tensed sentences like (3-5) and relative clauses like (6).

(3) Scarlet stabbed Mustard.
(4) Peacock saw me stab him.
(5) Scarlet didn’t stab Mustard.
(6) who she stabbed

Sentences can be analyzed as instructions for how to build monadic concepts—of a special sort that interface with other cognitive systems in a distinctive way—without invoking concepts of truth values. This requires a little more than M-junction and Θ-junction. But given this account of sentences, and a slight tweak on standard treatments of pronouns as in (4), relative clauses can also be accommodated. And while there is no avoiding appeal to something like quantification over assignments of value(s) to variables, such quantification—in a massively monadic me(n)tal language—is much less powerful than standard appeals to lambda abstraction.

This paves the way for a treatment of quantificational constructions like (7-9).

(7) Every guest stabbed him.
(8) She stabbed every guest.
(9) Most guests stabbed every spy.

As discussed in section three, a phrase like ‘every spy’ can be analyzed as an instruction for how to build a plural monadic concept of things—perhaps things as simple as ordered pairs—whose internal participants are the relevant spies. I assume that at least many quantificational expressions are grammatically displaced (or “raised”), and that such expressions recombine with sentential expressions that contain a trace of the displacement. But in my view, displaced quantifiers do not combine with expressions of type <e, t> that are formed by abstraction on expressions of type <t>: this standard approach to examples like (7-9) obscures important parallels with (1-4); and it effectively abandons the idea that meanings compose in accord with substantive constraints, replacing this idea with the much weaker claim that expressions have recursively specifiable semantic properties. Correlatively, I think my proposal—which might be
described as a neo-Medieval treatment of quantification—offers a better description of the much discussed “conservativity” constraint on determiners like ‘every’ and ‘most’. It also handles (10),

(10) every guest who stabbed him

which cannot be understood as a sentence with the meaning of (7). This unambiguity is puzzling if ‘every spy’ is of type \( \langle e, t \rangle \) and ‘who stabbed him’ is of type \( \langle e, t \rangle \). So not only can standard accounts of quantification be recoded. the proposed recoding is in some respects better.

1. Nouns, Verbs, and Simple Phrases

The generable expressions of a human I-language connect meanings with pronunciations. These expressions, formed by combining lexical items in certain ways, also exhibit syntactic properties that can be described independently of the associated semantic and phonological properties; though of course, the “autonomous” syntactic structures of human linguistic expressions constrain what the expressions can mean, and how they can be pronounced. Like many others, I think the meaning of a complex expression is tightly connected to the form of the expression, modulo the arbitrary character of lexical meanings. But in defending a conception of meaning, one must beware of positing implausibly powerful expression-generating capacities that are tailored to the conception. (Similarly, in defending a conception of syntax, one must beware of positing implausibly powerful operations of semantic composition.) So my aim is show that my proposed account of meaning is plausible given spare expression-generating capacities—and ideally, only the capacities assumed by any adequate conception of human languages; cp. Chomsky (1995, 2000a). As it happens, I think “minimalist” theories of syntax are independently plausible, on theoretical and empirical grounds; see, e.g., Hornstein (2001, 2009), Uriagereka et. al. (200x), Berwick et.al. (2011). But the main point here is that minimal expression-generating capacities suffice for my proposed semantics. Though if the net result is a descriptively adequate conception of syntax and semantics, with a spare syntax and semantics taking in each other’s wash, that would be a nontrivial argument for the component theories.\(^1\)

1.1 Building Simple Instructions

If human I-languages generate unboundedly many expressions from finitely many lexical items, then such I-languages are procedures that include at least one operation that somehow supports recursive combination of generable expressions. It must be possible to combine lexical items in a way that forms a “unit” that can combine with at least one lexical item to form another such unit. One can invent a language in which any two expressions can be “unified” in a way that is adequately represented with set formation: given any two expressions \( \Sigma \) and \( \Sigma' \), \( \text{UNIFY}(\Sigma, \Sigma') = \{\Sigma, \Sigma'\} \); so given any expression \( \Sigma'' \), \( \text{UNIFY}(\{\Sigma, \Sigma'\}, \Sigma'') = \{\{\Sigma, \Sigma'\}, \Sigma''\} \); etc. But obviously, human I-languages are not so permissive. Words exhibit categories that correspond to restrictions on what can combine with what; see, e.g., Baker (2003). Moreover, phrases are headed. Combining ‘brown’ with ‘cow’, for example, yields a phrase that is importantly more like ‘cow’ than ‘brown’. So I assume that at least some expressions/meanings are built up in way that goes beyond the “super spare” operation UNIFY.

\(^1\) See Hornstein and Pietroski (2009) for extended discussion. Correlatively, I understand the Minimalist Program broadly, not merely as an attempt to simplify extant conceptions of syntax. Like many authors—e.g., Fox (1999), Borer (2005), Jackendoff (2002), Ramchand (2008)—I assume that goal is to simplify syntax and semantics, as opposed to minimizing one at the cost of complicating the other.
Following Chomsky (1995, 2000a) and others, one might posit a basic asymmetric operation MERGE such that MERGE(\(\Sigma, \Sigma'\)) is either \{\{\Sigma, \Sigma'\}, \Sigma\} or \{\{\Sigma, \Sigma'\}, \Sigma'\}; where the singleton member of the set is the label of the expression formed by combining \(\Sigma\) with \(\Sigma'\). Everything I say in this chapter can be recast in these terms. But it will helpful to adopt a slightly different proposal that has some independent attractions. Instead of taking MERGE to be a basic operation, I adopt Hornstein’s (2009) idea that Chomsky’s posited operation reflects two simpler operations: a super spare operation like UNIFY, and an operation LABEL that can apply to results of the first operation.  

The idea is that MERGE(\(\Sigma, \Sigma'\)) = LABEL(UNIFY(\(\Sigma, \Sigma'\))) = LABEL({\{\Sigma, \Sigma'\}, \Sigma}); where this result is either \{\{\Sigma, \Sigma'\}, \Sigma\} or \{\{\Sigma, \Sigma'\}, \Sigma'\} depending on which constituent has the properties required to “head” the unit \{\Sigma, \Sigma'\}. If \(\Sigma\) is the noun ‘cow’, and \(\Sigma'\) is the adjective ‘brown’, then LABEL({\{\Sigma, \Sigma'\}, \Sigma'}) = \{\{‘cow’, ‘brown’\}, ‘cow’\}. But this set, identical with \{\{‘brown’, ‘cow’\}, ‘cow’\}, can also be described in a more familiar way: [\(\text{brown}_{A} \text{cow}_N\)]_N.

This leaves room for the hypothesis that lexical categories are themselves functional expressions that can combine with acategorial roots. Perhaps the noun ‘cow’ is a complex expression, \([\text{\(\nu\)} \text{cow}_N\]_N, whose label is literally its functional constituent. And perhaps as Chametsky (199x) argues, there are cases of labeled expressions combining with unlabeled adjuncts. For simplicity, I’ll assume that the indefinite article in ‘a cow’ and the plural marker in ‘cows’ are unlabeled atomic constituents of complex nouns, as in [a \(\text{cow}_N\)]_N and [\(\text{cow}_N\ s\]_N. But it won’t matter if these expressions are described as determiner phrases like [\(\text{a}_{D} \text{cow}_N\]_D or number phrases like [\(\text{cow}_N \text{s}_{D}\)]_D; cp. Munn (199x).

The expression [\(\text{brown}_{A} \text{cow}_N\]_N connects a certain meaning, \(\mu([\text{brown}_{A} \text{cow}_N\]_N), with a certain pronunciation, \(\pi([\text{brown}_{A} \text{cow}_N\]_N). I take the meaning of the complex expression to be a complex instruction, leaving it open how the structure of a meaning (or “semantic instruction”) is related to the structure of a pronunciation (or “phonological instruction”). Even two-word examples—e.g., ‘brown toast’, which is ambiguous as between the noun phrase [\(\text{brown}_{A} \text{toast}_N\]_N and the verb phrase [\(\text{brown}_{V} \text{toast}_N\]_V—illustrate potential complexities. But as this kind of notation highlights, a nonatomic expression has at least three potentially significant components: two constituents and a label. So we can ask how the meaning of a complex expression is related to the cognitive significance of lexical items, unifying, and labeling.

Unsurprisingly, I think the significance of unifying is conjunctive. My suspicion is that labeling lets humans use old forms of conjunction in new cognitively fruitful ways; see Hornstein and Pietroski (2009). But speculations about natural history aside, it is useful to think about phrases as products of the operations UNIFY and LABEL. As we’ll see, this provides a way of describing the conjunctive aspects of phrasal meanings while recognizing that often, the significance of phrasal syntax is not merely conjunctive.

\[^2\] Hornstein speaks of concatenation rather than unification. The difference is purely terminological. But since I am not concerned with pronunciation here, I want to avoid any suggestion of linear as opposed to compositional order.

\[^3\] Perhaps \(N\) is a device for fetching a functional monadic concept like INDEXABLE(\(\_\)), while \(V\) is a device for fetching a concept like TENSABLE(\(\_\)), thus allowing for a distinction between \(\mu(\text{chase}_V\})\) and \(\mu(\text{chase}_N\)\). I won’t pursue this suggestion, but simply raise it to note the resources available without appeal to Fregean typology; see Hornstein and Pietroski (2010), drawing on Marantz (1984), Halle and Marantz (1993), Baker (2003), and Borer (2005). See also Harley (200x, 200x).
1.2 Specifying Simple Instructions

Recall that $\textbf{M-join}(1, I')$ is a “macro” instruction that is executed by executing the two subinstructions and $\textbf{M}$-joining the results. Executing $\textbf{M-join}(1, I')$ yields a concept of the form $\Phi(\_)^\Psi(\_)$). Likewise, executing $\textbf{Θ-join}(1, I')$ yields a concept of the form $\exists[\Theta(\_, \_)^\Psi(\_)]$. A subinstruction can be of the form $\textbf{fetch}@L$; where executing such an instruction accesses a monadic or dyadic concept from the lexical address $L$. But a subinstruction can also be of the form $\textbf{M-join}(1, I')$ or $\textbf{Θ-join}(1, I')$. My proposal is that simple phrases are instructions for how to build monadic concepts in these simple ways. But the details depend on which concepts can be accessed with which lexical items, and exactly how labeling is related to meaning.

Some complex expressions at least approximate the classical/medieval ideal of a conjunctive monadic predicate. For example, $\text{[cow}_N\text{s}}_N$ arguably corresponds to the concept $\text{cow(\_)^{PLURAL(\_)}}$; where as described at the end of the last chapter, $\text{cow(\_)}$ is a number-neutral concept. So $\text{[cow}_N\text{s}}_N$ can perhaps be described as the instruction $\textbf{M-join}(\mu(\text{cow}_N), \mu(\text{s}))$. And if ‘cow that Fido chased’ has the structured displayed with $\text{[cow}_N[\text{that Fido chased}]]_N$—ignoring structure within the relative clause labeled ‘C’—then the meaning of this noun phrase can perhaps be described as $\textbf{M-join}(\mu(\text{cow}_N), \mu([\text{that Fido chased}]))$. Of course, this highlights the need for an account of how $\mu([\text{that Fido chased}])$ can be an instruction for how to build a mental predicate. But it also highlights the fact that the significance of combining a noun with an adjective is rarely if ever merely conjunctive.

If we idealize away from the fact that a brown cow is a cow that is brown for a cow, we might say that executing $\mu(\text{brown}_A)$ accesses or assembles a monadic concept $\text{brown(\_)},$ which can be $\textbf{M}$-joined with $\text{cow(\_)}. And perhaps $\mu([\text{brown}_A \text{cow}_N])$ can be correctly described as the instruction $\textbf{M-join}(\mu(\text{brown}_A), \mu(\text{cow}_N))$. Though executing $\mu(\text{brown}_A)$ might access or assemble a dyadic concept $\text{brown-for(\_, \_)}$; where this concept applies to one more things (e.g., a cow) and some things (e.g., the cows) iff the former is/are both among the latter and brown instances of the latter. So perhaps the syntax of ‘brown cow’ is more like $\text{[brown-one}_A \text{cow}_N; where the meaning of the complex adjective $\text{[brown-one}_A$ is an instruction for how to build a concept like $\exists[\text{brown-for(\_, \_)}^{\text{THE-COWS(\_)}}$. In which case, $\mu([\text{brown-one}_A \text{cow}_N])$ could be the instruction $\textbf{M-join}(\mu([\text{brown-one}_A), \mu(\text{cow}_N))$. But executing this instruction might yield a conjunctive but not merely conjunctive concept like $\exists[\text{brown-for(\_, \_)}^{\text{THE-COWS(\_)}}^{\text{cow(\_)}}$. Indeed, the concept formed by $\textbf{M}$-junction could have a constituent formed by $\textbf{Θ}$-junction.

Examples like ‘big ant’, ‘toy truck’, and ‘fake diamond’ are sometimes described as non-conjunctive. But this strikes me as misdescription. A big ant is still an ant that meets a further condition on its size. A fake diamond is a fake of a certain sort. And depending on the context, a toy truck is either a toy of a certain sort, or a truck of a certain sort. (See Higginbotham 198x.) So before concluding that any such example provides a special difficulty for the view presented here, we need to know what syntax is being presupposed and what alternative captures the meaning better given that syntax. If difficult cases raise questions about how to capture the not merely conjunctive aspects of meaning, then we may be better off trying to supplement an account of meaning that is conjunctive at its core, instead of starting with a non-conjunctive account and trying to capture all the apparently conjunctive aspects of meaning.
Given the history of semantics, it is tempting to describe the meaning of a verb phrase like ‘chase cows’ or ‘chase a cow’ in terms of saturation or function-application. But upon reflection, such phrases show signs of being conjunctive though not merely conjunctive. Events of chasing cows are chased, though of course, not chased that are also cows. And while the meaning of ‘chase a cow’ is presumably of the same type as the meaning of ‘chase Bessie’, we should not assume that ‘Bessie’ is of type <e>, given the considerations adduced in chapter five.

One can hypothesize that in ‘chase a cow’, ‘a cow’ is a quantificational expression that displaces, leaving a trace of type <e> that corresponds to the first saturater of a concept like CHASE(\_, \_) or PASSIVE-CHASE(\_, \_). But the availability of this hypothesis hardly shows that the meaning of ‘chase a cow’ is non-conjunctive. As discussed below, traces need not be described as expressions of type <e>, and ‘a’ need not be described as a quantifier. Moreover, just as a chase of a cow may be both a chase and one that was done by a dog, it may have been done to a cow. And even if there is no intuitive sense in which such a chase is “of a cow,” we can speak of the dog’s chase of the cow. If ‘chase a cow’ had the meaning that ‘a chase or a cow’ has, or the meaning of ‘eat a cow’, then we might well describe the significance of combining ‘chase’ with ‘a cow’ as non-conjunctive, much as the idiomatic significance of combining ‘kick’ with ‘the bucket’ is non-conjunctive. But in my view, ‘kick the bucket’ still has one meaning that is conjunctive though not merely conjunctive, and likewise for ‘chase a cow’.

Thinking about phrases as products of the operations UNIFY and LABEL suggests a way factoring the not merely conjunctive significance of combining a verb with a noun (phrase) into two components: a conjunctive operation invoked by UNIFY; and an “adaptive” operation—involving a thematic concept like PATIENT(\_, \_)—invoked by the formal relation that a verb bears to the noun it combines with and dominates. Again, the meaning of [chaseV cowsN]\_V is clearly not of the simple form M-join(μ(chaseV), μ(cowsN)), since executing μ(cowsN) yields a concept like COW(\_)^PLURAL(\_). But there is more than one potential alternative.

One option is to hypothesize, a la Kratzer, that executing μ(chaseV) accesses the dyadic PASSIVE-CHASE(\_, \_); cp. §2.4 of chapter five. On this view, μ(chaseV cowsN)\_ is the instruction Θ-join(μ(chaseV), μ(cowsN)), whose execution can yield a concept of events that are collectively chases of cows: 3[PASSIVE-CHASE(\_, \_)^\_][COW(\_)^PLURAL(\_)]. This treats verbs as relevantly like the prepositions in [aboveP cowsN]p. I assume that prepositions can be used to access dyadic concepts, and that the meaning of [aboveP cowsN]p is an instruction whose execution can yield a concept like 3[ABOVE(\_, \_)^\_][COW(\_)^PLURAL(\_)]. But this doesn’t require a system that explicitly specifies the meanings of prepositional phrases in terms of Θ-junction and the meanings of plural nouns in terms of M-junction. One might characterize phrasal meanings as “neutral” instructions of the form Join(I, I’); where such an instruction can be executed by performing either of the more specific instructions M-join(I, I’) or Θ-join(I, I’).

If μ(aboveP) can only be executed by fetching a dyadic concept, then the neutral instruction Join(μ(aboveP), μ(cowsN)) can only be executed by Θ-joining the results of executing the two subinstructions. Likewise, If μ(chaseV) can only be executed by fetching a dyadic concept, then Join[μ(chaseV), μ(cowsN)] can only be executed by Θ-joining the results of executing the two subinstructions. But if each subinstruction in Join[μ(cowsN), μ(s)] can only be executed by fetching a monadic concept, then the macro-instruction can only be executed by M-joining the results of executing μ(cowsN) and μ(s).
In the end, perhaps we will conclude that verbs do access dyadic concepts, and that the significance of combining a verb with direct object is like saturation to this extent: the result has a lower adicity than one of the inputs. But I don’t want to assume that verbs access dyadic concepts, rather than monadic concepts like CHASE(). Moreover, we eventually have to deal with ‘saw a dog chase cows’ and its constituent ‘a dog chase cows’, whose meaning is not an instruction of the form JOIN[µ(dogsN), µ([chaseV cowsN]V)]. So let me present another option, according to which µ([chaseV [a cowsN]N]V) is an instruction whose execution can yield a concept like CHASE( )³PATIENT( , )³COWS( ), which can be M-joined with ³AGENT( , )³DOG( ).

Suppose that if expressions with the labels L and L* can be combined—yielding an expression of the form [, … , L*]L—then the label-pair L/L* determines a concept like PATIENT( , ) or perhaps IDENTICAL( , ). The idea is that the labeling in an expression of the form [, … , L*]L can serve as instruction to access a third concept, which can be described as the L/L*-adapter, in addition the concepts accessed or assembled by executing the constituent instructions µ(, ) and µ(, ). Obviously, one wants to limit the number of such “adapter concepts.” But for simplicity, let’s suppose that each label-pair also serves as the address for some such concept, allowing for the possibility that a single concept has several addresses of the form ‘L/L*’. Then one can posit a general composition principle (11)

(11) µ([, … , L*]L) = JOIN(µ(, ), JOIN(fetch@L/L*, µ(, )))

that has consequences like (12).

(12) µ([chaseV, cowsN]V) = JOIN(µ(chaseV), JOIN(fetch@V/N, µ(cowsN)))

From this perspective, combining expressions can call for M-junction and Θ-junction. Suppose the V/N-adapter is the dyadic concept PATIENT( , ). Then the instruction JOIN(fetch@V/N, µ(cowsN)) can only be executed by Θ-joining PATIENT( , ) with the result of executing µ(cowsN), thereby creating a monadic concept like ³PATIENT( , )³COWS( ). So if µ(chaseV) can only be executed by fetching a monadic concept, then µ([chaseV, cowsN]V) must be executed via M-junction, yielding a concept like CHASE( )³PATIENT( , )³COWS( ).

Of course, (11) also implies (13).

(13) µ([aboveP, cowsN]P) = JOIN(µ(aboveP), JOIN(fetch@P/N, µ(cowsN)))

But this may not be a bad thing. Svenonius (2010, 201x) argues that at least in some languages, prepositions correspond to monadic concepts of states, and that the preposition/verb distinction is not so clear; see also Schein (2014). But (13) may be correct even if prepositions like aboveP fetch dyadic concepts like ABOVE( , ). For one could say that in languages where prepositions access dyadic concepts, the label-pair P/N determines an “empty” concept like IDENTICAL( , ) or SELF-IDENTICAL( ), so that executing JOIN(fetch@P/N, µ(cowsN)) yields a monadic concept like ³IDENTICAL( , )³COWS( ) or SELF-IDENTICAL( )³COWS( ) or some other concept that is analytically equivalent to COWS(). This allows for a kind of variation in how relationality is expressed—via certain lexical items, or via thematic adapters associated with certain labels.

Likewise, if verbs access dyadic concepts in some but not all human languages, then one might say that the label-pair P/N has no real (as opposed to merely formal) significance in some but not all human languages. I worry about introducing this kind of freedom and assuming that children somehow “figure out” how relationality is expressed in a particular language. But

---

4 This does not preclude the possibility that L = L*.
there is whatever variation there is. And my point here is simply that for better or worse, my proposal allows for meanings/instructions that call for M-junction, Θ-junction, or both. So for better or worse, my proposal allows for more than one conception of verbs and prepositions.

If the significance of combining expressions is ever merely conjunctive, positing at least one “empty” concept will be unavoidable given (11), which also implies (14).

\[
(14) \mu([\text{cow}_N \text{ that Fido chased}]_{\text{c}}_{\text{N}}) = \text{Join}(\mu(\text{cow}_N), \text{Join}(\text{fetch@N/C}_N, \mu([\text{that Fido chased}]_{\text{c}})))
\]

But again, one can say the label-pair N/C determines a concept like IDENTICAL( _, _). If the plural marker in ‘cows’ has no label, then (11) does not apply. And we can say that if at least one constituent of a complex expression is not labeled, the significance of the combination is mere M-junction, so that \(\mu([\text{cow}_N \text{ s}]_{\text{N}}) = \text{M-join}(\mu(\text{cow}_N), \mu(s))\). But if the plural marker has a label, as shown in (15), then perhaps the #/N-adapter is purely formal,

\[
(15) \mu([\text{cow}_{s#}]_{\theta}) = \text{Join}(\mu(\text{cow}_{s}), \text{Join}(\text{fetch@#N}, \mu(s)))
\]

Obviously, M-junction can be viewed as a special case of Θ-junction:

M-joining the concepts \(\Phi(\_\) and \(\Psi(\_\) is equivalent to M-joining \(\Phi(\_\) with the result of Θ-joining \(\Psi(\_\) with IDENTICAL( _, _). So one might think of labels as a way of introducing the possibility of mismatching labels, and thereby injecting some non-conjunctive significance into many but not all instances of grammatical combination. Imagine a mind that can M-join each of its monadic concepts with some others, but only in limited ways, given various typological restrictions. Merely adding a capacity to Θ-join some monadic concepts with some dyadic concepts might not have a large cognitive impact. But given a capacity to lexicalize concepts in a way that introduces neutral concepts that can be systematically M-joined/Θ-joined, then it might very useful to have a system that (i) labels the resulting lexical items, which can be used to fetch the neutral concepts, and (ii) supports recursive combination of labeled expressions in a way makes it possible to use label pairs like V/N as devices for accessing concepts like patient(_, _), and agent(_, _); see Hornstein (200x), Hornstein and Pietroski (2009).

### 1.3 Internals and Externals

However, a substantive V/N-adapter need not be the specific concept patient(_, _). This would imply that executing the meaning of any phrase of the form \([…V, …N]_v\) yields a concept of the form \(\Phi(\_\right)^{\mu[\text{patient}(\_\right)^{\mu[\text{ψ}(\_\right)]}\]. But one might think that executing \(\mu([\text{see}_v, \text{cows}_N]_v\) can yield \(\text{see}(\_)^{\mu[\text{theme}(\_\right)^{\text{cows}(\_\right]}\); where this concept can apply to events without implying that any cows were affected. And plausibly, \(\text{see}(\_)^{\mu[\text{patient}(\_\right)^{\text{cows}(\_\right]}\) only applies to events in which cows were affected. So if verbs access monadic concepts like see(_, _) and chase(_, _), then given the general composition principle (11),

\[
(11) \mu([\ldots, \ldots, \ldots]_l) = \text{Join}(\mu([\ldots]_l), \text{Join}(\text{fetch@\text{L}/L^*}, \mu([\ldots\text{L}\right]))
\]

it might be better to say that the V/N-adapter is the more formalistic concept internal(_, _); where events of chasing can have internal participants that are patients of those events, events of seeing can have internal participants that are themes of those events, the ordered pair \(\{2, 3\}\) — a.k.a. \(\{2, \{2, 3\}\}\) has 3 as its internal participant, and so on.

On this view, both \(\mu([\text{chase}_v, \text{cows}_N]_v\) and \(\mu([\text{see}_v, \text{cows}_N]_v\) access the concept internal(_, _). But \(\text{chase}(\_)^{\mu[\text{internal}(\_\right)^{\text{cows}(\_\right]}\) can still differ thematically from \(\text{see}(\_)^{\mu[\text{internal}(\_\right)^{\text{cows}(\_\right]}\) if chase(_, _) carries an implication that it only applies to
events whose internal participants are patients, while \textit{see}() does carry this implication. One can say that \textit{chase}() and \textit{see}() are neutral with regard to many aspects of conceptual typologies, but that \textit{chase}() and \textit{see}() are still typologically distinct in some respects; cp. Vendler’s (1965a) much discussed hierarchy.

A more pressing concern is that whatever one says about the V/N-adapter has to be compatible with the facts concerning \textit{transitive} clauses like (16).\footnote{I want to delay discussion of tense and complete sentences until section two.}

(16) a dog chase cows

As discussed in chapter five, (16) cannot be understood as (an instruction for how to build) a predicate of chases whose agents are cows and whose patient is a dog. But if executing \(\mu([[\text{chase}\_V, \text{cows}_N]\_V])\) can yield \textit{chase}()\(^\gamma\)\(\exists[\text{INTERNAL}(\_, \_)\^\gamma\text{cows}(\_)\text{]}\)\text{]}—with \textit{chase}() licensing replacement of \textit{INTERNAL}(\_, \_) with \textit{patient}(\_, \_)—then so far, so good. However, this view requires that (16) not have the simple structure [[\text{a dog}_N][\text{chase}\_V, \text{cows}_N]\_V]. Otherwise, executing the meaning of this expression would yield a concept like (17),

(17) \(\exists[\text{INTERNAL}(\_, \_), \gamma\text{dog}(\_), \gamma\text{chase}(\_), \gamma\text{cows}(\_)]\)

which applies to a chase only if (a) its internal participant is a dog, and (b) its internal participants are cows. And that won’t do. So if (16) has the simple structure, (11) is wrong.

(11) \(\mu([\ldots, L)^\gamma]\) = \text{Join}(\mu([\ldots, L^\gamma]); \mu([\ldots, L^\gamma]))

One alternative is to say that (16) has the structure [[\text{a dog}_N][\text{chase}_V, \text{cows}_N]\_V]; where numerals indicate the number of arguments that an expression can still combine with. But not only does this recapitulate talk of “verbal valence,” in a way at odds with the spirit of chapter five, it requires a more sophisticated conception of labeling. Another alternative, explored in Pietroski (2005a), would be to replace (11) with (11a) and (11b);

(11a) \(\mu([\ldots, L^{\text{atomic}}), \ldots, L^\gamma]) = \text{Join}(\mu([\ldots, L^\gamma]); \mu([\ldots, L^\gamma]))

(11b) \(\mu([\ldots, L^{\text{phrasal}}), \ldots, L^\gamma]) = \text{Join}(\mu([\ldots, L^\gamma]); \mu([\ldots, L^\gamma]))

where \textit{chase}_V counts as atomic instance of category V, and [[\text{chase}_V, \text{cows}_N]\_V counts as a phrasal instance of the same category V, while ‘int’ and ‘ext’ are addresses for the formalistic concepts \textit{internal}(\_, \_) and \textit{external}(\_, \_). Then executing \(\mu([[\text{a dog}_N][\text{chase}_V, \text{cows}_N]\_V])\) could yield (18), with \textit{chase}() licensing replacement of \textit{external}(\_, \_) with \textit{agent}(\_, \_).\footnote{A variation on this suggestion is that “cyclicity” plays a role in choosing adapters: within a given “phase” of instruction \textit{execution}, the innermost occurrence label ‘V’ accesses \textit{internal}(E, X); the outermost occurrence of ‘V’ accesses \textit{external}(E, X). If there are at most two grammatical arguments per phase/cycle/whatever, one might imagine a binary “switch” that gets “reset” to its initial state at the start of each cycle; cp. Boeckx (2008).}

(18) \(\exists[\text{external}(\_, \_), \gamma\text{dog}(\_), \gamma\text{chase}(\_), \gamma\text{cows}(\_)]\)

A third option is to say that (16) has the structure [[\text{a dog}_N][\text{chase}_V, \text{cows}_N]\_V]; where [[\text{chase}_V, \text{cows}_N]\_V combines with a covert “small verb” to form a phrase that can combine with [[\text{a dog}_N]. On this increasingly common view, \textit{chase}_V itself only combines with a single grammatical argument. But one need not assign any significance to the covert verb, which can be viewed as a functional item whose role is to support the second grammatical argument: it may access no concept, or least none that applies any independent restriction; and the v/V-adapter may be the concept \textit{identical}(\_, \_). Then since (11) implies (19),
(19) \( \mu(\_v \text{[chase}_v, \text{cows}_N]_v) = \)
\[ \text{Join}(\mu(\_v), \text{Join}(\text{fetch}@v/V, \mu([\text{chase}_v, \text{cows}_N]_v))) \]
executing \( \mu(\_v \text{[chase}_v, \text{cows}_N]_v) \) could yield a concept like (20),
\[ (20) \exists [\text{IDENTICAL}(_, _)^\text{CHASE}(\_)]\exists [\text{INTERNAL}(_, _)^\text{COWS}(\_)] \]
which is equivalent to \( \text{CHASE}(\_)^\text{CHASE}(\_)^\text{COWS}(\_) \). In this sense, the covert small verb might do nothing except provide a new label for \( \text{[chase}_v, \text{cows}_N]_v \). But this is still doing something if the \( v/N \)-adaptor is \( \text{EXTERNAL} \). For in that case, given that (11) implies (21),
\[ (21) \mu([[a \text{ dog}_N]_v \text{[chase}_v, \text{cows}_N]_v],_v) = \]
\[ \text{Join}(\mu([\text{chase}_v, \text{cows}_N]_v), \text{Join}(\text{fetch}@v/N, \mu([a \text{ dog}_N])) \]
executing \( [[a \text{ dog}_N]_v \text{[chase}_v, \text{cows}_N]_v] \). Can yield (18).

Given a genuinely causative construction like (22),
\[ (22) \text{a child break vase} \]
whose verb can appear in the intransitive construction (23),
\[ (23) \text{a vase break} \]
one might posit a more substantive “light verb” as in (22a).
\[ (22a) [[a \text{ child}_N]_v \text{[termi}_v, \text{break}_v, [\text{a vase}_N]_N]_v] \]
Suppose the covert verb \( \text{termi}_v \), accesses a concept that applies to \( [e1, e2] \) if \( e1 \) is a causal process \text{terminates in} event \( e2 \); where a process of breaking a vase can terminate in an event that is a vase breaking.\(^7\) Then executing \( \mu([\text{termi}_v, \text{break}_v, [\text{a vase}_N]_N]_v)_v \) might yield (23),
\[ (23) \exists [\text{TERMINATE}(\_)]\exists [\text{CHASE}(\_)]\exists [\text{INTERNAL}(_, _)^\text{COWS}(\_)] \]
to which \( \exists [\text{EXTERNAL}(_, _)^\text{CHILD}(\_)] \) can be \( \Theta \)-joined. So the meaning of (22) can be specified as \( \text{Join}(\mu([\text{child}_N]_N), \text{Join}(\mu([\text{termi}_v, \text{break}_v, [\text{a vase}_N]_N]_v)_v).) \)

With this in mind, let’s return to the idea that executing \( \mu([\text{chase}_v] \) accesses the dyadic concept \( \text{PASSIVE-CHASE}(\_). \) Initially, one might think this option—which lets one say that
\( \mu([\text{chase}_v, \text{cows}_N]_v) = \text{Join}(\mu([\text{chase}_v], \mu([\text{cows}_N]))— \)is simpler than any alternative according to which executing \( \mu([\text{chase}_v] \) accesses a monadic concept. But if executing \( \mu([\text{chase}_v, \text{cows}_N]_v) \) yields \( \exists [\text{PASSIVE-CHASE}(_, _)^\text{COWS}(\_)] \), then to accommodate (22) and (16),
\[ (16) \text{a dog chase cows} \]
one still needs some principle according to which an external argument like ‘dog’ can correspond to a conjunct like \( \exists [\text{EXTERNAL}(_, _)^\text{DOG}(\_)] \). One can, of course, assign meanings of a suitably high Fregean type to the covert verbs; perhaps (16) indicates the function (24),
\[ (24) \lambda. E \cdot T = \exists x[\text{EXTERNAL}(E, x)^\text{DOG}(x)] \exists x[\text{PASSIVE-CHASE}(E, x)^\text{COWS}(x)] \]
and the covert verb indicates a function that maps \( \lambda. E \cdot T = \text{DOG}(x) \) to (24). Or one can simply posit a composition principle for external arguments that has the same effect; see Krakter (1996) appeal on “event identification.” But this is not simpler that uniform appeal to \( \Theta \)-junction and \( \Theta \)-junction.

Let me note, however, that appealing to \( \Theta \)-junction and thematic concepts effectively permits “conversion” of concepts of dogs and cows into concepts of events in which dogs and cows participate. As noted in chapter one, this is some respects like more familiar appeals to type-shifting in a fundamentally “saturationist” account of semantic composition. One might say
\[ ^7 \text{See Pietroski (2003b), drawing on a great deal of prior work, for presentation in a minimalist setting.} \]
that in the absence of labeling, the meaning of a phrase with two constituents would be an
instruction that is executed by executing the two subinstructions and saturating one result with
the other. But suppose that μ(\text{brown}_A) and μ(\text{cow}_N) can only be executed by accessing the
concepts \(\lambda x. T \equiv \text{brown}(x)\) and \(\lambda x. T \equiv \text{cow}(x)\). Then one might say that the labels in the phrase
[\text{brown}_A \text{cow}_N]_N matter, and that the label-pair \text{N/A} is an instruction to access the higher-order
concept \(\lambda \Phi. \lambda \Psi. \lambda x. T = \Psi(x) \& \Phi(x) = T\). This concept of type \(<<e, t>, <e, t>, <e, t>>\) can
be saturated by \(\lambda x. T = \text{brown}(x)\) to yield \(\lambda \Phi. \lambda x. T = \text{brown}(x) \& \Phi(x) = T\); where this latter
concept can be saturated by \(\lambda x. T = \text{cow}(x)\) to yield \(\lambda x. T = \text{brown}(x) \& \text{cow}(x) = T\). So
abstracting away from the brown-for-a-cow aspect of μ([brown_A cow_N]_N), one might specify this
meaning as follows: \textbf{Saturate(Saturate(fetch@ 'N/A', μ(brown_A)), μ(cow_N))}.

If only because this analysis of [brown_A cow_N]_N invokes a concept at Level Three of the
Fregean hierarchy, in addition to a concept of conjunction, it is not simpler than an analysis of
[chase_v cow_N]_V that invokes Θ-junction. But my proposal is analogous, in that \text{cow}(\_\_\_) can be
“Θ-adapted” to form the concept \(\exists\text{INTERNAL}(\_\_\_, \_\_\_, \text{cow}(\_\_\_))\). So while μ([chase_v cow_N]_V) is
executed by M-joining two monadic concepts, I don’t want to disguise the appeal to Θ-junction.

That said, it may be possible to impose a further and rather severe restriction on (11).
\(\mu([\ldots, \_\_\_, L^*]) = \text{Join}(\mu([\ldots, \_\_\_]), \text{Join}(\text{fetch@ 'L/L*'}, \mu([\ldots, \_\_\_])))\)
Perhaps the only “Θ-adapters” are \text{INTERNAL}(\_\_\_, \_\_) and \text{EXTERNAL}(\_\_\_, \_\_) and \text{IDENTICAL}(\_\_\_, \_\_).

1.4 Adverbials and Perceptuals

Given that event analyses were designed to treat adverbs and prepositional phrases as conjuncts
in complex predicates of events, it is not surprising that such expressions can be viewed as
instructions for how to build concepts of events. If the meaning of (16)
\(\text{a dog chase cows}\)
can be an instruction whose execution yields (18),
\(\exists[\text{EXTERNAL}(\_\_\_, \_\_)^\text{DOG}(\_\_\_)]^\text{CHASE}(\_\_)^\exists[\text{INTERNAL}(\_\_\_, \_\_)^\text{COWS}(\_\_\_)]\)
then the meaning of (25) can be an instruction whose execution yields (26).
\(\text{a dog chase a cow today}\)
\(\exists[\text{EXTERNAL}(\_\_\_, \_\_)^\text{DOG}(\_\_\_)]^\text{CHASE}(\_\_)^\exists[\text{INTERNAL}(\_\_\_, \_\_)^\text{COW}(\_\_\_)]^\text{TODAY}(\_\_)\)

As noted in §1.2 above, the meaning of \([\text{above}_p \text{cows}_N]_p\) can be an whose execution
yields a concept like \(\exists[\text{ABOVE}(\_\_\_, \_\_)^\text{COWS}(\_\_\_)]\). Likewise, executing μ([\text{with}_p \text{a knife}_N]_p) can
yield \(\exists[\text{INSTRUMENT}(\_\_\_, \_\_)^\text{KNIFE}(\_\_\_)]\). There need not be a single concept that the polysemous
preposition ‘with’ accesses; cp. ‘France’ and ‘book’. But if prepositions can be used to access
thematic adapters, then clauses like (1) can be used to build concepts like (1a);
\(\text{a spy stab a colonel with a knife}\)
\(\exists[\text{EXTERNAL}(\_\_\_, \_\_)^\text{SPY}(\_\_\_)]^\text{STAB}(\_\_)^\exists[\text{INTERNAL}(\_\_\_, \_\_)^\text{COLONEL}(\_\_\_)]^\exists[\text{INSTRUMENT}(\_\_\_, \_\_)^\text{KNIFE}(\_\_\_)\]
where \(\text{STAB}(\_\_)\) licenses replacement of the formalistic \text{EXTERNAL}(\_\_\_, \_\_) and \text{INTERNAL}(\_\_\_, \_\_) with
the more specific thematic concepts \text{AGENT}(\_\_\_, \_\_) and \text{PATIENT}(\_\_\_, \_\_). And if proper nouns are like
phrases headed by common nouns—in accessing monadic concepts, albeit concepts that apply to
only one thing—then clauses like (27) can be used to build concepts like (27a).
\(\text{Scarlet stab Mustard}\)
\(\exists[\text{EXTERNAL}(\_\_\_, \_\_)^\text{SCARLET}(\_\_\_)]^\text{STAB}(\_\_)^\exists[\text{INTERNAL}(\_\_\_, \_\_)^\text{MUSTARD}(\_\_\_)]\)
At this point, accommodating perceptual idioms like (2) is easy.

(2) see Scarlet stab Mustard

If (27) is the complement of ‘see’ in (2), then (2) is relevantly like (28), which implies (29),

(2) see a stab of Mustard that was done by Scarlet
(29) see a stab

The meaning of (29) can be described as an instruction whose execution can yield (29a).

(29a) see(\_)^∃[internal(\_,\_)^stab(\_)]

Likewise, the meaning of (2) can be described as an instruction whose execution can yield (2a).

(2a) see(\_)^∃[internal(\_,\_)^∃[external(\_,\_)^scarlet(\_)^stab(\_)^∃[internal(\_,\_)^mustard(\_)]]]

This concept applies to events of seeing whose internal participants (or themes) are events of Scarlet stabbing Mustard.

2. Tense, Sentences, and Relative Clauses

It still needs showing that the proposed account of untensed clauses can be extended to accommodate sentences like (3-6) and relatival clauses like (6).

(3) Scarlet stabbed Mustard.
(4) Peacock saw me stab him.
(5) Scarlet didn’t stab Mustard.
(6) who she stabbed

This is not entirely trivial. But it does not require appeal to truth values, or to concepts that cannot be inputs to the basic composition operations of M-junction and Θ-junction.

2.1 Adding Tense

Adding a tense marker to (27)

(27) Scarlet stab Mustard

can be viewed as instruction to build a temporal concept and M-join it with a concept like (27a).

(27a) ∃[external(\_,\_)^scarlet(\_)^stab(\_)^∃[internal(\_,\_)^mustard(\_)]]

Following Hornstein (198x) and others, I suspect that a concept of events like past(\_) is best viewed as a complex Reichenbach-style concept, perhaps ∃[before(\_,\_)^speech-time(\_)]; where the concept speech-time(\_) is context sensitive concept and perhaps itself of the form ∃[Θ(\_,\_)^Φ(\_)]. In my view, Hornstein’s analysis highlights some reasons for wanting a fundamentally conjunctive (as opposed to saturationist) account of human tenses. But in any case, his account lends itself to expression in terms of M-junction and Θ-junction.

Let’s assume that (27) is a verb phrase of some kind, and describe it simply as follows:

[scarlet stab mustard]v; where the subscript may reflect both ‘stab’ and a covert small verb. And lets assume that a tense morpheme can be combined with (27) to form a tense phrase like

[edt [scarlet stab mustard]v]t; where for purposes of pronunciation, the head of this phrase can be “lowered” onto the verb ‘stab’ (see Pollack 199x). Then the meaning of the past tense morpheme can be described in much the same way as the meaning of the temporal adverb in (30)

(30) Scarlet stab Mustard yesterday

---

8 Grammatical aspect is its own complex topic that I will not try to address here. But see Hacquard (2009, 2010). Her account of is encoded in traditional type-theoretic terms. But as Hacquard’s proposed logical forms make clear, her event-based conception of aspect (and its relation to tense) is underlyingly conjunctive, and her leading ideas can be captured terms of Θ-junction.
or perhaps the adverbial phrase in (31).

(31) Scarlet stab Mustard before midnight

That is, \( \mu([\mathit{ed}_{T} \ [\mathit{Scarlet stab Mustard}]_{V}]_{T}) \) can be described as the following instruction:

\[ \text{Join}(\mu(-\mathit{ed}_{T}), \text{Join}(\mathit{fetch}('@T/V', \mu([\mathit{Scarlet stab Mustard}]_{V}))) \].

If executing \( \mu(-\mathit{ed}_{T}) \) yields a monadic concept—say, \( \exists[\mathit{before}(_{-}, _)^{\mathit{speech-time}(_{})}] \)—there is no need for a T/V-adapter. In which case, \( \mu([-\mathit{ed}_{T} \ [\mathit{Scarlet stab Mustard}]_{V}]_{T}) \) has the same effect as the formally simpler instruction \( \mathbf{M-join}(\mu(-\mathit{ed}_{T}), \mu([\mathit{Scarlet stab Mustard}]_{V})) \).

This captures the subtle ambiguity of (32),

(32) Peacock hear Scarlet shoot Green yesterday

which is structurally homophonous as between (32') and (32''),

(32') \([\text{Peacock hear } [[\mathit{Scarlet shoot Green} ]_{V} \text{ yesterday}]_{A}]_{V}\)

(32'') \([[[\text{Peacock hear } \ [\mathit{Scarlet shoot Green}]_{V} \text{ yesterday}]_{A}]_{V}\)

either of which can be a constituent of a question like (33); see Higginbotham (198x).

(33) Did Peacock hear Scarlet shoot Green today?

If the shot occurred just before midnight, but Peacock heard it just after midnight, then the answer to (33) might be positive on the (32')-reading but negative on the (33')-reading; cp. the discussion in \( \S \times x \) of chapter three. But adding a tense morpheme to (32), on either reading, would unambiguously add the temporal restriction to \( \mathit{hear}(_{-}) \) rather than \( \mathit{shoot}(_{-}) \).

However, if adding a tense morpheme corresponds to just another case of M-junction, then executing \( \mu([-\mathit{ed}_{T} \ [\mathit{Scarlet stab Mustard}]_{V}]_{T}) \) yields a concept like (34).

(34) \[ \exists[\mathit{before}(_{-}, _)^{\mathit{speech-time}(_{})}]^{\wedge} \exists[\mathit{external}(_{-}, _)^{\mathit{scarlet}(_{-})}]^{\wedge} \exists[\mathit{internal}(_{-}, _)^{\mathit{mustard}(_{-})}]\]

And while \( \mu([-\mathit{ed}_{T} \ [\mathit{Scarlet stab Mustard}]_{V}]_{T}) \) can be paired with the pronunciation of (34),

(34) Scarlet stabbed Mustard

which can be viewed as tensed clause, this clause differs somehow from the “complete sentence” (3), whose period corresponds to something like an existential closure of the open slot in (34).

(3) Scarlet stabbed Mustard.

This can make it tempting to describe \( \mu([\mathit{Scarlet stab Mustard}]_{V}) \) and \( \mu(-\mathit{ed}_{T}) \) as meanings of types \( <e, t> \) and \( <<e, t>, t> \). For example, one might specify \( \mu(-\mathit{ed}_{T}) \) as the function \( \lambda \Phi. T = \exists \Phi(e) & \mathit{past}(e) \). But this stipulates that a tense morpheme, unlike the adverb in (30),

(30) Scarlet stab Mustard yesterday

does two jobs: add temporal information, like a \textit{conjunctive adjunct}; and close a variable, like a \textit{quantification argument}. Alternatively, one might say that (3) is the result of adding a covert existential quantifier—of type \( <<e, t>, t> \)—to \( [-\mathit{ed}_{T} \ [\mathit{Scarlet stab Mustard}]_{V}]_{T} \). But this raises the question of why such a quantifier cannot be pronounced, with (35) having the meaning of (3).

(35) Some event Scarlet stabbed Mustard.

I don’t doubt that some such account can be sustained. But there is another way of describing the difference between (3) and (30), according to which even (3) is an instruction for how to build a monadic concept.

\subsection*{2.2 T-Concepts}

A monadic concept of events is neither true or false. One can say that (34)
(34) \[\exists \text{BEFORE}(\_, \_)^\text{SPEECH-TIME}(\_)^\wedge \exists \text{EXTERNAL}(\_, \_)^\text{SCARLET}(\_)^\wedge \text{STAB}(\_)^\wedge \exists \text{INTERNAL}(\_, \_)^\text{MUSTARD}(\_)]\] is satisfied by an assignment \(A\) of value(s) to variables iff \(A\) assigns to the “zeroth” variable—where by stipulation, this is the variable associated with the open slot in (34)—one or more things that are both events of Scarlet stabbing Mustard and things that occurred prior to whatever time \(A\) assigns to the variable associated with ‘now’. Correlatively, one can say that relative to any given time, (34) is true of stabbings of Mustard that were done by Scarlet prior to that time. But (34) is not itself true or false, not even relative to a time. However, it doesn’t follow that if a concept has a satisfaction condition, it cannot be true or false. On the contrary, Tarski (1933) showed how to describe truth in terms of satisfaction. As discussed in chapter two, the closed sentences of a Tarskian language are not expressions of a special type \(<t>\). They are, however, sentences that are satisfied by each assignment or by none. Likewise, (34) can be “closed” in a way that yields another monadic concept, albeit one that applies to everything or nothing.

Consider a pair of operators, \(\uparrow\) and \(\downarrow\), that create monadic concepts from monadic concepts; where for any one or more things, \(\uparrow \Phi(\_\) applies to them iff \(\Phi(\_)\) applies to one or more things, and \(\downarrow \Phi(\_)\) applies to them iff \(\Phi(\_)\) applies to nothing. More briefly, and without worrying about number neutrality: for each entity, \(\uparrow \Phi(\_)\) applies to it iff \(\Phi(\_)\) applies to something; and \(\downarrow \Phi(\_)\) applies to it iff \(\Phi(\_)\) applies to nothing. One can think of \(\uparrow\) and \(\downarrow\) as polarizing operators—cp. §2.2 of chapter one—that convert any monadic concept into a concept of everything or nothing, perhaps akin to \(\text{EXIST}(\_)\) and \(\neg \text{EXIST}(\_)\). For example, given any entity, \(\uparrow \text{COW}(\_)\) applies to it iff \(\text{COW}(\_)\) applies to something; so \(\uparrow \text{COW}(\_)\) applies to you, and likewise to me, iff there is a cow. By contrast, \(\downarrow \text{COW}(\_)\) applies to you (and to me) iff nothing is a cow. And for each thing, either \(\uparrow \text{COW}(\_)\) or \(\downarrow \text{COW}(\_)\) applies to it—since each thing is either such that there is a cow, or such that there is no cow. Correlatively, nothing is such that both \(\uparrow \text{COW}(\_)\) and \(\downarrow \text{COW}(\_)\) apply to it. Hence, nothing is such that \(\uparrow \text{COW}(\_)^\wedge \downarrow \text{COW}(\_)\) applies to it.

Given a suitable metalanguage, we can say: \(\forall x \{\uparrow \Phi(x) \equiv \exists y[\Phi(y)]\}\); and similarly, \(\forall x \{\downarrow \Phi(x) \equiv \neg \exists y[\Phi(y)]\}\). But the idea is not that ‘\(\uparrow\)’ and ‘\(\downarrow\)’ are abbreviations for Tarski-style existential closure and its negation. For example, ‘\(\uparrow \text{BETWEEN}(x, y, z)\)’ is gibberish, as is ‘\(\downarrow \text{AGENT}(e, x)\)’. The idea is rather that certain human linguistic meanings, perhaps associated with tense and/or negation, invoke “closure” operations that convert a monadic concept (say, of events) into a concept of all or none. So let’s say that any concept of the form \(\uparrow \Phi(\_)\) or \(\downarrow \Phi(\_)\) is a \(T\)-concept, with ‘\(T\)’ connoting Tarski, truth, and totality.

Note that for any concept \(\Phi(\_)\) and any entity \(e\), \(\uparrow \downarrow \Phi(\_\) applies to \(e\) iff \(\downarrow \downarrow \Phi(\_\) does, since each of these concepts applies to \(e\) iff \(\uparrow \Phi(\_)\) does—i.e., iff \(\Phi(\_)\) applies to something. Likewise, \(\uparrow \downarrow \Phi(\_\) applies to \(e\) iff \(\downarrow \downarrow \Phi(\_\) does, since each of these concepts applies to \(e\) iff \(\downarrow \Phi(\_)\) does—i.e., iff \(\Phi(\_)\) applies to nothing. And while \(\uparrow [\Phi(\_)^\wedge \Psi(\_)\]\) applies to \(e\) iff something falls under the conjunctive concept \(\Phi(\_)^\wedge \Psi(\_)\), which applies to \(e\) iff \(e\) falls under both conjuncts, \(\uparrow \Phi(\_)^\wedge \Psi(\_)\]\) applies to \(e\) iff (e is such that) something falls \(\Phi(\_)\) and something falls

\(^9\) Or if you prefer, for any one or more things: [\(\uparrow \Phi(\_)\)](\_) applies to them iff \(\Phi(\_)\) applies to one or more things; and [\(\downarrow \Phi(\_)\)](\_) applies to them iff \(\Phi(\_)\) applies to nothing. But omitting the extra slot and brackets turns out to be at least as perspicuous. For these purposes, let me ignore mass nouns and assume countable values of variables.
under $\Psi(\_\,)$. Thus, $\uparrow[BROWN(\_\,) \land COW(\_\,)]$ is more restrictive than $\downarrow[BROWN(\_\,)^{\uparrow} \land COW(\_\,)]$, much as the more familiar (36) implies (37), but not conversely.

(36) $\exists x[BROWN(x) \land COW(x)]$

(37) $\exists x[BROWN(x)] \land \exists x[COW(x)]$.

Correlatively, $\downarrow[BROWN(\_\,)^{\uparrow} \land COW(\_\,)]$ applies to $\_\,$ iff nothing is both brown and a cow, while $\downarrow[BROWN(\_\,)^{\uparrow} \land COW(\_\,)]$ applies to $\_\,$ iff nothing is brown and nothing is a cow. So $\downarrow[BROWN(\_\,)^{\uparrow} \land COW(\_\,)]$ is more restrictive than $\downarrow[BROWN(\_\,)^{\uparrow} \land COW(\_\,)]$. And $\downarrow COW(\_\,)^\uparrow$ is more restrictive than $\downarrow[BROWN(\_\,)^{\uparrow} \land COW(\_\,)]$; again, cp. the earlier discussion of “medieval polarity.”

Note that when the concepts conjoined are themselves T-concepts, which apply to all or none, “closing up” has no effect. If $P$ and $Q$ are T-concepts, and so each is of the form $\uparrow \Phi(\_\,)$ or $\downarrow \Phi(\_\,)$, then $\uparrow[P \land Q]$ is logically equivalent to $P \land Q$: $\uparrow[P \land Q]$ applies to $\_\,$ iff something/everything falls under both $P$ and $Q$; $P \land Q$ applies to $\_\,$ iff e/everything falls under both $P$ and $Q$. By contrast, $\downarrow[P \land Q]$ applies to $\_\,$ iff: nothing falls under both $P$ and $Q$; i.e., nothing is such that both $P$ and $Q$ are empty; i.e., something falls under $P$ or something falls under $Q$. So propositional disjunction can be characterized, a la de Morgan, given T-concepts.

So we should pause before assuming that human I-languages generate expressions of the Fregean type $\langle t \rangle$, as opposed to expressions that can be used to build T-concepts, which can bear an intimate relation to existential thoughts of type $\langle t \rangle$. Let’s abbreviate (34) as (34a),

(34) $\exists \,[BEFORE(\_\_, \_)^{\uparrow} \land SPEECH\,-\,TIME(\_\,)]^{\uparrow}$

(34a) $\exists \,[EXTERNAL(\_\_, \_)^{\uparrow} \land SCARLET(\_\,)]^{\uparrow} \land \exists \,[INTERNAL(\_\_, \_)^{\uparrow} \land MUSTARD(\_\,)]$

and assume that this is the concept built by executing $\mu([-\,ed_T [\text{Scarlet stab Mustard}]_V]_T)$. If (3)

(3) Scarlet stabbed Mustard.

is the result of combining the homophonous tensed clause with a silent polarity head as in (3’)

(3’) $[+Pol [-\,ed_T [\text{Scarlet stab Mustard}]_V]_T]_{Pol}$

then (3) can analyzed as the following instruction: $\text{CloseUp}(\mu([-\,ed_T [\text{Scarlet stab Mustard}]_V]_T))$; where this instruction is executed by executing the embedded instruction, thereby creating (34a), and applying the operator $\uparrow$ to form the corresponding T-concept, (3a).

(3a) $\uparrow \text{PAST\,-\,STAB\,-\,OF\,MUSTARD\,-\,BY\,SCARLET(\_\,)}$

This clearly invokes a non-conjunctive (polarizing) operation. So principle (11)

(11) $\mu([-L_{\ldots \_ \ldots} \ldots L^*_{\ldots \_ \ldots}]_L) = \text{Join}(\mu([-L_{\ldots \_ \ldots} \ldots L^*_{\ldots \_ \ldots}]), \text{Join}(\text{fetch}^{\downarrow}_{\_ \ldots} \downarrow L^*_{\ldots \_ \ldots}, \mu([-L_{\ldots \_ \ldots} \ldots L^*_{\ldots \_ \ldots}]))$

must be limited to tensed clauses (including aspect in the tense system for these purposes). From this perspective, the polarity head marks a “new phase” of composition in which an instruction that conforms to (11) is “topped off” in certain ways, perhaps in ways that correspond to preparing a constructed monadic concept for certain uses; see Lohndal and Pietroski (2011) for related discussion drawing on Chomsky (2005). But even if the meaning of (3) is an instruction to perform a closure operation on a certain concept (obtained by executing a tensed instruction), the net result is still monadic concept. And I don’t think that this analysis of (3) is any less simple than positing a covert quantifier of type $\langle <e, t \rangle, t \rangle$. If (5) has the structure shown in (5’),

(5) Scarlet didn’t stab Mustard.

(5’) $[\text{notPol} [-\,ed_T [\text{Scarlet stab Mustard}]_V]_T]_{Pol}$

with “sentential” negation characterized as an overt negative polarity item, then (5) can analyzed
as the following instruction: \texttt{CloseDown}(\mu([-ed_T [Scarlet stab Mustard]_V]_T)); where this instruction is executed by executing the embedded instruction, thereby creating (34a), and applying the operator to form the corresponding \texttt{T-concept}, (5a).

\begin{equation}
(5a) \Downarrow_{\text{PAST-STAB-OF-MUSTARD-BY-SCARLET(\_)}}
\end{equation}

A mind that can form monadic concepts like (3a) and (5a) might also be able to form the corresponding \textit{thoughts}, (3b) and (5b),

\begin{align*}
(3b) \exists e \&\text{PAST-STAB-OF-MUSTARD-BY-SCARLET(e)} \\
(5b) \sim \exists e \&\text{PAST-STAB-OF-MUSTARD-BY-SCARLET(e)}
\end{align*}

or perhaps subject-predicate analogs of these existential generalizations. Such a mind might treat (3a) and (5a) as cognitively equivalent to (3b) and (5b), at least for many purposes. But while “pre-linguistic” and “post-linguistic” cognition presumably traffics in complete thoughts, at least often, human I-languages may \textit{interface with} such cognition via formally monadic \texttt{T-concepts}. Indeed, one can view (3a) and (5a) as formal variants of (3b) and (5b), which cannot themselves be outputs of (or inputs to) \texttt{M-junction} or \texttt{Θ-junction}.

\subsection*{2.3 Abstraction}

Given the operators $\uparrow$ and $\downarrow$, a mind can convert a concept that applies to some but not all things into a \texttt{T-concept} that applies to all or none. But this does not yet explain the converse capacity, much less the capacity to form complex concepts like \texttt{WHO-SCARLET-STABBED(_)}. And this is arguably the most interesting respect in which human thought is recursive. Reformulating a point from chapter one: given \texttt{P\textasciitilde Q}, the capacity to form \texttt{P\textasciitilde[P\textasciitilde Q]} is not that impressive; and in important respects, the recursion exhibited by ‘who saw a colonel that Scarlet stabbed’ is more interesting than that exhibited by ‘thinks that Peacock said that Scarlet stabbed Mustard’. Embedding one sentence in another is a good trick. Using a sentence to create a concept that applies to some but not all things is a great trick. Clearly, this requires more than mere conjunction. Though as Tarski (1933) showed, the requisite machinery is not especially complicated: it involves a distinctive operation—in effect, quantification over ways of assigning values to variables—but not saturation. So if as I suspect, appeal to the Tarskian operation is unavoidable in theory of meaning for a human I-language, then we need ask whether such a theory of should \textit{also} appeal to a Frege-Church operation along with a hierarchy of types.

Let’s assume that humans have \textit{conceptual indices} that can be used as devices for temporarily tracking salient things (cp. Pylyshyn [2007]). Some of these indices may be singular context-sensitive denoters of type $<e>$. But suppose that human I-languages allow for \textit{grammatical indices} like those indicated with the numerical subscripts in [she$_1$ stab him$_2$]$_V$ and [them$_3$ stab it$_4$]$_V$. And suppose these indices can be used to access context-sensitive mental predicates like \texttt{FIRST(_)} and \texttt{SECOND(_)}. The idea is that in any context, \texttt{FIRST(_)} applies to whatever is tracked with the first index used in that context, and likewise for \texttt{SECOND(_)}.

If it helps, we can say that \texttt{SECOND(_)} is satisfied by an assignment $\mathcal{A}$—of values to potential variables—iff the one or more things that $\mathcal{A}$ assigns to the (zeroth) variable associated with the open slot are the one or more things that $\mathcal{A}$ assigns to the second potential variable; likewise for \texttt{FIRST(_)}. Then one can view the meanings of [-ed$_T$ [she$_1$ stab him$_2$]$_V$]; and [-ed$_T$ [them$_3$ stab it$_4$]$_V$]; as instructions for how to build concepts like (38) and (39).
on whether or not any assignment quantification over assignments of values $t$ Fregean type. But for any given variable type, lambda abstraction is equivalent to Tarski

$$\mu([\text{me stab him}_1]_\forall)$$ as an instruction for how to build a concept like (40).

So sentences like (4) can also be accommodated, given the proposed account of clauses like (2).

(4) Peacock saw me stab him.

(2) see Scarlet stab Mustard

Note that if grammatical indices are analyzed as expressions of type $<e>$, then one needs to account for the contrasting predicative aspects of ‘she’ and ‘him’ without supposing that these pronouns are results of combining an expression of type $<e, t>$ with an index. But let’s ignore gender for simplicity, and abbreviate (38-40) as (38a-40a), suppressing thematic structure.

$$\begin{align*}
(38a) & \text{PAST-2-STAB-1}(\_)
(39a) & \text{PAST-3-STAB-4}(\_)
(40a) & \text{PAST-2-SPEAKER-STAB-1}(\_)
\end{align*}$$

Corresponding to (38a), there are two T-concepts. The concept $\uparrow\text{PAST-2-STAB-1}(\_)$ applies to everything or nothing, relative to assignment $\mathbb{A}$, depending on whether or not: prior to whatever time $\mathbb{A}$ assigns to the variable associated with ‘now’, there were some stabbings of whatever $\mathbb{A}$ assigns to the second index done by whatever $\mathbb{A}$ assigns to the first index. The concept $\downarrow\text{PAST-2-STAB-1}(\_)$ applies to nothing iff $\uparrow\text{PAST-2-STAB-1}(\_)$ applies to everything. Let’s abbreviate these T-concepts, in bold, as follows: $\text{2-STABBED-1}(\_); \text{NOT-2-STABBED-1}(\_)$. Likewise, let’s abbreviate the T-concepts corresponding to (39a) and (39b) as follows: $\text{3-STABBED-4}(\_); \text{NOT-3-STABBED-4}(\_); \text{SPEAKER-STABBED-1}(\_); \text{NOT-SPEAKER-STABBED-1}(\_)$.

These are analogs of conditional assignment-relative specifications of truth values like ‘T if there was a stabbing by $\mathbb{A}(1)$ of $\mathbb{A}(1)$, and $\bot$ otherwise’. But T-concepts apply to everything or to nothing; they do not specify truth values. One can say that this is notational variation. But then appealing to truth values and expressions of type $<e>$ is a potentially misleading notational variant of appeal to instructions for how to build T-concepts. In any case, a mind that can form T-concepts like $\text{2-STABBED-1}(\_)$—perhaps via introduced concepts that can be M-joined and $\Theta$-joined to form monadic concepts that can be “closed” with the operator $\uparrow$—does not need the full power of Church-style lambda abstraction to form concepts like WHO-SCARLET-STABBED(\_).

As discussed in chapter two, the invented language CPL generates denoters of every Fregean type. But for any given variable type, lambda abstraction is equivalent to Tarski-style quantification over assignments of values to variables of that type. So for example, relative to any assignment $\mathbb{A}$: the complex denoter $\lambda x. T \equiv \text{SPY}(x)$ denotes the smallest function $F$ of type $<e, t>$ such that for each assignment $\mathbb{A}^*$ such that $\mathbb{A}^*=x \mathbb{A}$, $F$ maps $\mathbb{A}^*(x)$ to $T$ or $\bot$, depending on whether or not $\mathbb{A}$ satisfies the open sentence $\text{SPY}(x)$; and likewise, the complex denoter $\lambda x. T \equiv \text{ABOVE}(\text{SCARLET}, x) \& \text{ABOVE}(x, x')$’ denotes the smallest function $F$ of type $<e, t>$ such
that for each assignment $\mathcal{A}^*$ such that $\mathcal{A}^*\equiv_\mathcal{A}$, F maps $\mathcal{A}^*(x)$ to $T$ or \bot, depending on whether or not $\mathcal{A}^*$ satisfies the open sentence $\textsc{above(Scarlet, X)} \& \textsc{above(X, X')}$.

As I said in chapter two, lambda abstraction can be viewed as a way of unleashing the power of quantifying over sequence variants within a language that extends the Tarskian language PL in a Fregean way. But PL already allowed for open sentences with arbitrarily many variables. So given the Fregean hierarchy of types, CPL generates concepts all of those types, starting with the more than two million types at Level Four or lower, the more than $5 \times 10^{12}$ types at Level Five, etc. One might hope that sparer resources would suffice to let a mind form concepts like $\textsc{who-Scarlet-stabbed(\_)}$ and $\textsc{who-stabbed-mustard(\_)}$. And indeed, a mind that can form T-concepts like $\textsc{2-stabbed-1(\_)}$ only needs the basic Tarskian trick—of quantifying over assignment variants—to use a relative clause like (6)

(6) who she stabbed

as an instruction whose execution can yield a concept like who-Scarlet-stabbed(\_).

Given any index $i$ and T-concept $P$, which can evaluated relative to any assignment $\mathcal{A}$, let $\textsc{Tarski\{i, P\}}$ be a “semantic” concept whose content can be specified with (41);

(41) $\exists \mathcal{A}^*: \mathcal{A}^*\equiv_\mathcal{A}\{\text{assigns}(\mathcal{A}^*, \_, i) \& \text{satisfies}(\mathcal{A}^*, P)\}$

where $\text{assigns}(\mathcal{A}^*, \_, i)$ applies to whatever $\mathcal{A}^*$ assigns to the $i$th index, and ‘$\mathcal{A}^*\equiv_\mathcal{A}\mathcal{A}$’ means that $\mathcal{A}^*$ differs from $\mathcal{A}$ at most with regard to what it assigns to the $i$th index. To be sure, any natural concept of satisfaction is likely to differ from Tarski’s; and $\textsc{Tarski\{i, P\}}$ is posited as a natural concept, presumably less intellectualized than (41). But my suggestion is that we can use a T-concept like $\textsc{2-stabbed-1(\_)}$ to form a concept that applies to one or more things (relative to an assignment of values to indices) iff making them the values of the first index (and holding everything else constant) satisfies the T-concept. One can think of this as a restricted number-neutral version of lambda abstraction. But the idea is not that we apply a general capacity for lambda abstraction to (6). The idea is human I-languages support Tarski-style abstraction on T-concepts that contain indices corresponding to concepts like $\textsc{first(\_)}$ and $\textsc{second(\_)}$.

Relative to any assignment $\mathcal{A}$, and ignoring the tense for simplicity, the concept $\textsc{Tarski\{1, 2-stabbed-1(\_)}$ applies to one or more things iff they were stabbed by whatever $\mathcal{A}$ assigns to ‘2’. We can specify this application condition more formally with (42);

(42) $\exists \mathcal{A}^*: \mathcal{A}^*\equiv_1\mathcal{A}\{\text{assigns}(\mathcal{A}^*, \_, 1) \& \text{satisfies}(\mathcal{A}^*, \textsc{2-stabbed-1(\_)}\}$

if $\mathcal{A}^*\equiv_1\mathcal{A}$, $\mathcal{A}^*$ assigns whatever $\mathcal{A}$ assigns to ‘2’. And we can abbreviate (42) with (42a).

(42a) $\lambda 1. \textsc{2-stabbed-1}$

But we shouldn’t let this abbreviation make us think that $\textsc{Tarski\{1, 2-stabbed-1(\_)}$ is a concept of a function that maps entities to truth values. Similarly, $\textsc{Tarski\{2, 2-stabbed-1(\_)}$ applies to one or more things iff they stabbed whatever $\mathcal{A}$ assigns to ‘1’. We can specify this application condition more formally with (43), which we can abbreviate with (43a).

(43) $\exists \mathcal{A}^*: \mathcal{A}^*\equiv_2\mathcal{A}\{\text{assigns}(\mathcal{A}^*, \_, 2) \& \text{satisfies}(\mathcal{A}^*, \textsc{2-stabbed-1(\_)}\}$

(43a) $\lambda 2. \textsc{2-stabbed-1}$

But abbreviation doesn’t show that $\textsc{Tarski\{2, 2-stabbed-1(\_)}$ maps entities to truth values, much less that (44) denotes a truth value relative to each assignment of values to indices in (44).
Likewise, executing ABSTRACT forming a T
For simplicity, let’s assume that the displaced position is associated with a covert complementizer that takes a sentential expression is a displaced analog of ‘him
my view, the question is whether we posit
But avoid not that this kind of abstraction can be reduced to anything else. On the contrary, I see no way to avoid positing a capacity for construction in minds that can understand relative clauses like (6).
(6) who she stabbed
But in offering plausible theories of meaning for human I-languages, I think we also need to posit—one way or another—the more mundane operations of M-junction and Θ-junction respect. Relative to a particular assignment \( A \), the \( 2\text{-STABBED-1(\_)} \) may apply to nothing, while \( \text{TARSKI}\{1, 2\text{-STABBED-1(\_)}\} \) applies to many things; see Salmon (2006). Suppose that whatever \( A \) assigns to ‘2’, it/they stabbed many things, but not whatever \( A \) assigns to ‘1’. Then relative to \( A \): \( 2\text{-STABBED-1(\_)} \) is false of each thing, and in that sense false; yet \( \text{TARSKI}\{1, 2\text{-STABBED-1(\_)}\} \) is true of many things, and so \( \upmodels \text{TARSKI}\{1, 2\text{-STABBED-1(\_)}\} \) is true of each thing. And if whatever \( A \) assigns to ‘2’ stabbed nothing, then \( 2\text{-STABBED-1(\_)} \) and \( \upmodels \text{TARSKI}\{1, 2\text{-STABBED-1(\_)}\} \) are false relative to \( A \). Like it or not, this “non-truth-functional” composition is available to a mind that is equipped to perform the Tarski trick. And my claim is not that this kind of abstraction can be reduced to anything else. On the contrary, I see no way to avoid positing a capacity for construction in minds that can understand relative clauses like (6).
(44) She stabbed him.
If \( 2\text{-STABBED-1(\_)} \) doesn’t denote a truth value, \( \text{TARSKI}\{2, 2\text{-STABBED-1(\_)}\} \) does not denote a function of type \( \langle e, t \rangle \). And one can’t establish that \( 2\text{-STABBED-1(\_)} \) denotes a truth value by insisting that \( \text{TARSKI}\{2, 2\text{-STABBED-1(\_)}\} \) denotes a function of type \( \langle e, t \rangle \). But I agree that using \( 2\text{-STABBED-1(\_)} \) to build a “relative clause concept,” with an application condition like that of (42) or (43), is an operation considerably more sophisticated than conjunction.
Indeed, this kind of concept construction violates certain compositionality constraints that M-junction and Θ-junction respect. So in my view, the question is whether we also need Fregean typology and an operation of saturation.

I haven’t yet specified the meaning of (6). But let’s assume, standardly, that the ‘wh’-expression is a displaced analog of ‘him2’ in the tensed phrase \([-edT [\text{she}_1 \text{ stab him}_2]_V]_T\); where the displaced position is associated with a covert complementizer that takes a sentential complement. If sentencehood is associated with a polarity head, as suggested above, then (6) has a grammatical structure along the lines of (6):

\[
(6') [\text{who}_1 [\_C [\_P [\_T [\text{she}_2 \text{ stabv who}_1]_V]_P]_T]_P]_C
\]

For simplicity, let’s assume that the covert complementizer is semantically null—that it is there simple to support the extraction of ‘who’—and that \( \mu([\_C [\_P [\_T [\text{she}_2 \text{ stabv who}_1]_V]_P]_T]_P]_C) \) just is the meaning of the polarity phrase. For any index \( i \), let an instruction of the form

\[
\text{ABSTRACT}(i, \mu([\_P]_P)) \text{ be executed by executing the instruction } \mu([\_P]_P) \text{—thereby forming a T-concept like } 2\text{-STABBED-1(\_)} \text{—and forming the corresponding concept of the form } \text{TARSKI}\{i, P\}. \text{ Then the meaning of (6) can be specified as the following instruction:}
\]

\[
\text{ABSTRACT}(1, \mu([\_P [\_T [\text{she}_2 \text{ stabv who}_1]_V]_P]_P]).
\]

Executing this meaning can yield \( \text{TARSKI}\{1, 2\text{-STABBED-1(\_)}\} \); cp. (42/42a).

\[
(42) \exists \text{A*}: \text{A*} =_1 \text{A} \{\text{ASSIGNS(A*, \_1) & SATISFIES(A*, 2\text{-STABBED-1(\_)})}
\]

\[
(42a) \lambda 1.2\text{-STABBED-1}
\]

Likewise, executing the meaning of (45) can yield \( \text{TARSKI}\{2, 2\text{-STABBED-1(\_)}\} \); cp. (43/43a).

(45) who stabbed him

\[
(43) \exists \text{A*}: \text{A*} =_2 \text{A} \{\text{ASSIGNS(A*, \_2) & SATISFIES(A*, 2\text{-STABBED-1(\_)})}
\]
corresponding to the formation of concepts and relative clauses (along with indices) to build corresponding concepts of the form Tarski \{i, P\}.  

3. Quantification

The meanings of quantificational constructions like (7) and (8) can now be accommodated.

(7) Every guest stabbed him.
(8) She stabbed every guest.

I assume that quantifiers like ‘every spy’ displace as in the structures shown below.

\[
\begin{align*}
[[\text{every guest}]_{Q2} & [+_{Pol} [\text{[every guest]}_2 \text{stabbled}_v \text{him}_1]_{T_{Pol}}]_Q \\
[[\text{every guest}]_{Q1} & [+_{Pol} [\text{she}_2 \text{stabbled}_v [\text{every guest}]_1]_{T_{Pol}}]_Q
\end{align*}
\]

For present purposes, the details regarding labeling do not matter. The idea is simply that in (7), the indexed quantifier \[[\text{every guest}]_{Q2}\] combines with a polarity phrase that can be used to build the T-concept 2-STABBED-I(\_); while in (8), the indexed quantifier \[[\text{every guest}]_{Q1}\] combines with a polarity phrase that can be used to build the same T-concept. But the idea is not that the indexed quantifiers combine with expressions that have the meanings of relative clauses like (45) who stabbed him.

On the contrary, I want to explain why (10) cannot be understood as synonymous with (7).

(10) every guest who stabbed him

---

10 As Heim and Kratzer’s (1998) system highlights, even if one appeals to saturation as a composition operation, one still needs to posit Tarskian abstraction—often encoded as lambda abstraction—as a distinct operation, even given a Fregean typology. Suppose we treat indices and traces as constituents, as in [2 [she1 chasedv t2s]], with the embedded sentence as an expression of type \(<e, t>\) and the larger expression as of type \(<e, t>\). From an I-language perspective, one can say (modulo tense and gender) that relative to any assignment \(\mathcal{A}\): the concept formed by executing \([\text{she1 chased t2s}]\) denotes truth iff whatever \(\mathcal{A}\) assigns to 1 chased whatever \(\mathcal{A}\) assigns to 2; and correlatively, the concept formed by executing \([2 [\text{she1 chasedv t2s}]\) applies to entity e iff whatever \(\mathcal{A}\) assigns to 1 chased e. But the idea *isn’t* and can’t be that the index denotes a function-in-extension of type \(<t, <e, t>>\), which maps the truth value of \([\text{she, chased t2s}]\) onto a function of type \(<e, t>\). Rather, ‘2’ has to indicate a hypothesized (syncategorematic) instruction to convert a representation of one sort into a representation of another sort. (Kobele [200x] shows how to pretend otherwise, for certain purposes, by positing a domain that includes assignments as entities in the basic domain. But in my view, this highlights the point.)

Heim and Kratzer’s third composition rule, in addition to rules for saturation and M-junction, makes this vivid. Their idea is that the higher copy of the lower index triggers quantification over assignment variants:

\[
[[2^{\lambda} [\text{she1 chasedv t2s}]]]^{\lambda} = \lambda x. \; T \; \text{iff} \; \exists \mathcal{A} \; \forall^* x = \mathcal{A}(2) \; \& \; [[\text{she1 chasedv t2s}]]^{\lambda^*} = T
\]

This has the desired result, taking the lambda-expression to be a theorist’s representation of the hypothesized concept obtained in two stages: execute the sentential instruction, obtaining a concept that is doubly sequence-sensitive, and modify the resulting concept as directed by the index ‘2’. One can remain agnostic about the detailed forms of the concepts constructed. And from an E-language perspective, one can take the lambda-expression to be (only) a theorist’s representation of the hypothesized satisfaction condition; cp. Kobele (200x). But from an I-language perspective, the goal is to say how competent speakers represent the alleged satisfaction condition. And while theorists can abbreviate—as in \([[2^{\lambda} [\text{she1 chasedv t2s}]]\]}^\lambda = \lambda x. \text{CHASE}(1, x)—we should remember that the corresponding psychological hypothesis presupposes some version of the Tarski trick.
My suggestion is that in (7), [every guest]_Q2 combines with a polarity phrase, not a relative clause. And while polarity phrases are used to build monadic concepts, of a special sort, the formal difference matters in ways that are mirrored by the meanings of quantifiers like ‘every’.

3.1 Quantifiers as Plural Predicates

The leading idea is simple: a quantificational determiner like ‘every’ (‘some’, ‘no’, ‘most’, etc.) accesses a number neutral concept of ordered pairs; where by stipulation, the ordered pair ⟨α, β⟩ has β as its internal participant and α as its external participants. More specifically, let’s say that every(_)_ applies to some ordered pairs iff: every one of their internal participants is one of their external participants; or put another way, (all of) their internals are among their externals. Likewise, most/some/no(_)_ applies to some ordered pairs iff most/some/none of their internals are among their externals.

This is analogous to saying that quantificational expressions are true of certain ordered pairs of sets. But every(_)_ applies, plurally, to the following six ordered pairs: ⟨1, 1⟩; ⟨2, 1⟩; ⟨3, 3⟩; ⟨4, 3⟩; ⟨5, 5⟩; ⟨6, 5⟩. Each of the three internal participants—1, 3, and 5—is one of the external participants. Likewise, every(_)_ applies, plurally, to the boundless many ordered pairs ⟨x, y⟩ such that x is a positive integer and y is a positive odd integer. But every(_)_ does not apply to ⟨{x: x is a positive integer}, {x: x is a positive odd integer}⟩. Nor does every(_)_ apply to this ordered pair of sets and the first six ordered pair of numbers, since one of internal participants—{x: x is a positive odd integer}— is not one of the external participants. But every(_)_ does apply, plurally, to the following three ordered pairs: ⟨1, 5⟩; ⟨3, 3⟩; ⟨5, 1⟩.

Let’s say that for any concept Φ(_), the concept MAX:Φ( _) applies to some things iff they are (all and only) the things to which Φ(_)_ applies. Put another way, MAX:Φ( _) applies to some things, the Xs, iff any one more things, the Ys, that Φ( _) applies to are among the Xs. Then the three-conjunct concept (46) applies to some ordered pairs iff:

\[
(46) \text{every}(\_){}^n \exists[\text{internal}(\_, \_){}^n \text{max:guest}(\_){}^n] \\
\exists[\text{external}(\_, \_){}^n \text{max:arrived}(\_)]
\]

each of their internal participants is one of their external participants; their internal participants are the guests; and their internal participants are the things that arrived. More briefly, (46) applies to some ordered pairs iff each of the guests is one of the things that arrived. Or still more briefly, (46) applies to some ordered pairs iff every guest arrived.

So using the operator † to “close up” (46) yields a T-concept that applies to everything or nothing, depending on whether or not every guest arrived. This should already make it clear that we have the resources needed to specify the meanings of (7) and (8),

(7) Every guest stabbed him.
(8) She stabbed every guest.

at least if they have the structures shown in (7)’ and (8)’.

(7)’ [+Pol [[everyQ guest]_N]_Q2 [+Pol [[every]_2 stabbed]_V him]_T]_Pol]_Q]_Pol
(8)’ [+Pol [[everyV guest]_N]_Q1 [+Pol [she2 stabbed]_V [every]_1]_T]_Pol]_Q]_Pol

What remains are some compositional details. In spelling these out, let’s assume that the concept internal(_ _, _) is the Q/N-adapter, while the concept external(_ _, _) is the Q/Pol-adapter. This corresponds to the familiar idea that a lexical quantifier combines with a noun to form a unit that
3.2 Building the Plural Predicates
If principle (11) applied to all labeled expressions,
\begin{equation}
(11) \mu([\ldots L_1, \ldots L^*]) = \text{Join}(\mu([\ldots L_1]), \text{Join}(\text{fetch}@\text{L/L}^*, \mu([\ldots L^*])))
\end{equation}
then \(\mu([\text{every}_Q, \text{guests}_N])\) would be the instruction (47),
\begin{equation}
(47) \text{Join}(\mu([\text{every}_Q]), \text{Join}(\text{fetch}@\text{Q/N}, \mu(\text{guests}_N)))
\end{equation}
whose execution would yield a concept like \(\text{EVERY}(\_)^\exists[\text{INTERNAL}(\_ \_)^\text{GUEST}(\_)]\). That is close to what we want. But we ‘every guest’ does not have the meaning of ‘every one of some guests’, it has the meaning of ‘every one of the guests’. One possible response is to posit a covert maximizing operator in the ‘every guest’ itself. But if we already need to say that (11) does not apply “above tense,” then another option is to posit the slight variation on (11) shown in (48);
\begin{equation}
(48) \mu([\ldots Q, \ldots N]) = \text{Join}(\mu([\ldots Q]), \text{Join}(\text{fetch}@\text{Q/N}, \text{Maximize}:\mu([\ldots N])))
\end{equation}
where for any instruction \(I\), the macro instruction \text{Maximize:} \(I\) is executed by executing \(I\)—thereby forming a monadic concept \(\Phi(\_)-\)and forming the corresponding concept \(\text{MAX:} \Phi(\_)-\). Then as desired, executing \(\mu([\text{every}_Q, \text{guests}_N])\) would yield a concept like (49).
\begin{equation}
(49) \text{EVERY}(\_)^\exists[\text{INTERNAL}(\_ \_)^\text{MAX:GUEST}(\_)]
\end{equation}
Let’s continue to assume that executing \(\mu([+\text{Pol} [[\text{every}_\text{guest}_2 \text{stabbed}_V \text{him}_1]]_T]_\text{Pol}])\) yields the concept \(2-\text{STABBED-1}(\_)-\); where this T-concept applies to everything or nothing, relative to any assignment \(A\), iff A(2) stabbed A(1). We have seen how \(2-\text{STABBED-1}(\_)-\) can be a constituent of the concept \(\text{TARSKI}\{2, 2-\text{STABBED-1}(\_)-\}\) as \(\lambda 2.2-\text{STABBED-1}-\) which applies, relative to any assignment \(A\), to one or more things iff they stabbed A(1). So relative to \(A\), concept (50) applies to things whose externals are the things that stabbed A(1).
\begin{equation}
(50) \exists[\text{EXTERNAL}(\_ \_)^\text{MAX:} \text{TARSKI}\{2, 2-\text{STABBED-1}(\_)-\}]
\end{equation}
And concept (50) can be built by executing instruction (51),
\begin{equation}
(51) \text{Join}(\text{fetch}@\text{Q/Pol} ,
\text{Maximize:} \text{ABSTRACT}(2, \mu([+\text{Pol} [[\text{every}_\text{guest}_2 \text{stabbed}_V \text{him}_1]]_T]_\text{Pol}]))
\end{equation}
So rule (52) will do the job.
\begin{equation}
(52) \mu([\ldots Q], \ldots \text{Pol}]) = 
\text{Join}(\mu([\ldots Q]), \text{Join}(\text{fetch}@\text{Q/Pol} ,
\text{Maximize:} \text{ABSTRACT}(i, \mu([\ldots \text{Pol}])))
\end{equation}
Then the meaning of the full Q-phrase in (7') can be specified as the following instruction:
\begin{equation}
(7') [+\text{Pol} [[\text{every}_Q \text{guest}_N_2]_T]_\text{Pol} [+\text{Pol} [[\text{every}_\text{guest}_2 \text{stabbed}_V \text{him}_1]]_T]_\text{Pol}]
\end{equation}
join the meaning of [everyQ guestN]Q2 with the meaning specified in (51). Executing this instruction can yield the M-junction of (49) with (50); and the \(\vdash\)-closure of this M-junction applies to every or nothing, relative to A, depending on whether or not every guest stabbed A(1). Likewise, the meaning of the full Q-phrase in (8') can be specified as the following instruction:
\begin{equation}
(8') [+\text{Pol} [[\text{every}_Q \text{guests}_N]_Q1]_T]_\text{Pol} [+\text{Pol} [[\text{she}_2 \text{stabbed}_V \text{every}_\text{guest}]]_T]_\text{Pol}]
\end{equation}
join the meaning of [everyQ guestN]Q1 with the meaning specified in (53).
\begin{equation}
(53) \text{Join}(\text{fetch}@\text{Q/Pol} ,
\text{Maximize:} \text{ABSTRACT}(1, \mu([+\text{Pol} [[\text{she}_2 \text{stabbed}_V \text{every}_\text{guest}]]]]_T]_\text{Pol})))
\end{equation}
Executing this instruction can yield the M-junction of (49) with (54);
\begin{equation}
(54) \exists[\text{EXTERNAL}(\_ \_)^\text{MAX:} \text{TARSKI}\{1, 2-\text{STABBED-1}(\_)-\}]
\end{equation}
and the ↑-closure of this M-junction applies to every or nothing, relative to 〈A, depending on whether or not 〈A(2) stabbed every guest.

If (9) has the structure shown in (9'),

(9) Most guests stabbed every spy. (9') 〈Pol [[mostQ guestsN]Q2] <> 
[+Pol [[everyQ spyN]Q1] <> 

then (9) presents no special difficulty. Executing the meaning of (9) yields a T-concept that applies to everything or nothing, depending on whether or not there are some ordered pairs that meet both the [mostQ guestsN]Q2 condition—i.e., the internal participants are the guests, and most of them are also external participants—and the following condition: their external participants are the things that (distributively) stabbed every spy; i.e., their external participants are the things such that for every spy, they (each) stabbed that spy. Spelling this condition out, in terms of quantification over assignments, so that it doesn’t matter (modulo tense) which assignment (9) is relativized to—makes the condition look complicated. But the meaning isn’t complicated. Or at least it is not more complicated than the meaning posited by more familiar treatments of (9) that appeal to saturation, Fregean typology, and rules for converting assignment-related expressions of type <t> into assignment-related expressions of type <e, t>; cp. note 10.

From this perspective, the compositional difference between [everyQ guestN]Q2 and a verb phrase like [seeV guestN]V is small. The quantificational phrase adds a maximality condition to the internal argument. Of course, one would like to reduce (48)

(48) 〈μ([…Ω, … γ]Ω) = Join(μ([…Ω], Join(fetch@Q/N, Maximize:μ([…γ]))

to a generalization of (11) that does apply “above tense” and a small twist regarding quantifiers,

(11) 〈μ([…L, … L*]L) = Join(μ([…L], Join(fetch@’L/L*, μ([…L*]))

(48) is not too ugly even as it stands. The proposal is not to invoke an entirely new principle of composition for quantifiers. On the contrary, the idea is to retain the idea that [everyQ guestN]Q2 and [seeV guestN]V are almost identical compositionally. To be sure, specifying the meaning of [everyQ guestN]Q2 as in (55) already builds the relevant maximization in via lambda abstraction.

(55) 〈λΦ.λΠ.Π = {x:Φ(x) = T} ⊆ {x:Π(x) = T} (λx.Π = guest(x))

Function (55), identical to 〈λΨ.Π = {x:guest(x)} ⊆ {x:Ψ(x) = T}, maps a function F of type <e, t> to T iff the extension of F includes every one of the guests. But precisely because lambda abstraction already includes the relevant kind of maximization, specifying the meaning of [seeV guestN]V in terms 〈λx.Π = guest(x) requires something extra—e.g., positing a covert existential quantifier, to get the meaning of ‘see some guests’ rather than ‘see (all) the guests’.

This doesn’t show that my proposal, which locates the extra twist in the quantificational phrase, is correct. But there is no a priori reason for thinking that it is better to locate the extra twist in the “bare plural” verb phrase.11 And I see no reason for insisting that the indefinite article in [seeV a guestN]V is an existential quantifier; see Higginbotham (1987). The difference

between \([\text{every}_Q \text{guest}_N]_{Q2}\) and \([\text{see}_V \text{guest}_N]_V\) has to be encoded somehow. One option is to encode it in a way that builds maximalization in via lambda abstraction, and de-maximalizes for certain cases. But another option is to adopt a principle like (11), which assigns weaker (un-maximalized) interpretations by default, and then add maximalization as needed—e.g., for strong quantifiers like ‘every’, which end up raising above tense.

Principle (52) is a little more \textit{ad hoc}.

\[(52) \mu([\ldots Q_i \ldots]_Q) = \text{Join}(\text{fetch}@'Q/\text{Pol}', \text{Maximize:ABSTRACT}(i, \mu(\ldots\text{Pol})))\]

Compared with (48), it introduces the further operation of abstraction. The idea is that when \([\text{every}_Q \text{guest}_N]_{Q2}\) combines a polarized phrase—which becomes the external argument of the raised quantifier \text{every}_Q—the \textit{index} is also relevant to the meaning, because the index of the quantifier tracks which aspect of polarized phrase is relevant to the abstraction. And because the index matters in this way, (52) is as syncategorematic as any other plausible account of quantification direct objects; see Jacobson (199x, 200x) and note 10 above.

So while (52) looks a little uglier than (48), which is not \textit{quite} as simple as (11), I think (48) and (52) have the virtue of making the relevant assumptions about \([\ldots Q \ldots N]_{QI} \ldots N_{\text{Pol}}]_{QI}\) explicit: the arguments of a quantifier are subject to a kind of maximization; the external argument is also subject to a kind of abstraction, dependent on the index. Moreover, the apparent complexity of (52) allows for a certain kind of simplification.

\textbf{3.2 Restriction and Conservativity}

Let’s return to (7).

\textit{(7)} Every guest stabbed him.

\textit{(7′)} \([\text{[every}_Q \text{guest}_N]_{Q2} [\text{[every}_V \text{guest}_2] \text{stabb}_V \text{him}_1]_{\text{Pol}}]_{Q}\)

In terms of this example, we don’t need to specify the relevant external participants as \textit{all the things} that stabbed \(\mathcal{A}(1)\). For purposes of capturing the thought built/expressed via (7), it is enough (and perhaps better) to specify the external participants to as the internal participants that stabbed \(\mathcal{A}(1)—\text{i.e., all the guests} that stabbed \(\mathcal{A}(1)\). If poor Colonel Mustard was also stabbed by some non-guests, that is irrelevant. This kind of point is often stressed when discussing the \textit{conservativity} of determiners; see Barwise and Cooper (1981), Higginbotham and May (1981).

Sentence (7) seems synonymous with (56).

\textit{(56)} Every guest is a guest who stabbed him.

This is interesting, especially given that (7) is not synonymous with (10),

\textit{(10)} every guest who stabbed him

which cannot be understood as a complete sentence. Specifying the meaning of \([\text{every}_Q \text{guest}_N]_{Q2}\) as in (55), simplified as (55a),

\textit{(55)} \(\lambda \Phi. \lambda \Psi. T \equiv \{x: \Phi(x) = T\} \subseteq \{x: \Psi(x) = T\}\)

\textit{(55a)} \(\lambda \Psi. T \equiv \{x: \text{GUEST}(x)\} \subseteq \{x: \Psi(x) = T\}\)

captures the equivalence of (7) and (56): the set of guests is a (perhaps improper) subset of the set of things that stabbed Mustard if the set of guests is a (perhaps improper) subset of the set of guests that stabbed Mustard. But one wants to know why (10) fails to have an equivalent reading—in addition to its actual reading in which the relative clause restricts “guest”—and why the “inverted” function specified in (57) is a meaning of any known quantificational determiner.

23
Characterizing the meanings of (7), (10), and (56) truth-theoretically seems like a way of missing what is important about these examples, especially if one describes them all in terms of a fundamentally “saturationist” conception of composition. But whatever one says about the difference between (7) and (10), one wants it to fit with what one says about why there is no analog of everyQ that indicates function (57).

Put another way, if (58) is a possible meaning for a lexical item of a human I-language,

\[(58) \lambda \Phi \lambda \Psi. T \equiv \{x : \Phi(x) = T\} \subseteq \{x : \Psi(x) = T\}\]

one wants to know why (57) is not attested. We can imagine a quantificational determiner gronkQ that is like everyQ, except that (59) has the meaning of (60).

(59) Gronk cow is brown.

(60) Every brown thing is a cow.

The absence of gronkQ is especially striking given (61),

(61) Only cows are brown.

which means roughly that all brown things are cows. It turns out to be independently plausible that ‘only’ is a focus operator, and not a quantifier that indicates function (57). But it seems unlikely that children allow for the possibility of such a quantifier, yet always end up concluding that there is no such quantifier.

Likewise, if (58) is a possible meaning for a lexical quantifier, one wants to know why (62) and (63) are not attested, especially given quantifiers like mostQ; see note x of chapter one.

\[(62) \lambda \Phi \lambda \Psi. T \equiv \text{IDENTICAL}\{\{x : \Phi(x) = T\}\} \subseteq \{x : \Psi(x) = T\}\]

\[(63) \lambda \Phi \lambda \Psi. T \equiv \text{EQUINUMEROUS}\{\{x : \Phi(x) = T\}\} \subseteq \{x : \Psi(x) = T\}\]

Why don’t we have lexical items identQ and equiQ such that (64) and (65)

(64) Ident cows are brown.

(65) Equi cows are brown.

have the meanings of (64a) and (65a), respectively?

(64a) The cows are the brown things.

(65a) The cows are equinumerous with the brown things.

The general constraint is often described in terms of Barwise and Cooper’s (1981) notion of conservativity, defined for set-theoretic relations like inclusion: \(\mathbf{R}(\_ , \_ )\) is a conservative relation iff: \(\forall s \forall t [\mathbf{R}(s , t) = \mathbf{R}(s \cap t , s)]\); cp. Higginbotham and May’s (1981) notion of intersectivity. One can go on to define the corresponding functions of type \(<et , <et , t>\) as

\[12\] See, e.g., Herburger (2000). Note that ‘only’ can be inserted anywhere in ‘He said that she likes him’.
conservative, and say that no determiners express nonconservative relations. But however the 
 constraint on sentence meanings is initially formulated, one need not adopt a lexicalist 
 conception of it. That is, one need not posit a filter that somehow excludes otherwise available 
 lexical meanings like (62) and (63). I have suggested that executing the meaning of a quantifier 
 like everyQ accesses a concept like EVERY(), which applies to ordered pairs such that (all of) 
 their internal participants are among their external participants. If meanings are described in 
 these terms, one wants to know why there is no analogous though “inverted” quantifier gronkQ 
 that accesses a concept of ordered pairs whose external participants are among their internal 
 participants. But there may be an answer. 

It is often observed that expressions like [everyQ guestN]Q are understood as restricted 
 quantifiers. So one might stipulate that all expressions of the form …Q access concepts of 
 ordered pairs whose external participants are among (or selected from) their internal participants. 
 Then one could say that meaning everyQ accesses the concept like IDENT(), which applies to 
 ordered pairs such that their external participants are (i.e., are all of) their internal participants. 
 Given a global restriction to concepts of ordered pairs whose externals are selected from their 
 internals, adding the specific restriction that the selection be total—or put another way, taking 
 the maximal permitted selection—is equivalent to imposing the restriction that the internals be 
 among the externals. If the guests are among the things that stabbed Mustard, then the guests are 
 the guests who stabbed Mustard. That is one way of getting at the equivalence of (7) and (56). 
 But this way of describing the equivalence employs rather than eschews a notion of identity. 

(7) Every guest stabbed him.

(56) Every guest is a guest who stabbed him.

In one sense, it is a stipulation to say that lexical quantifiers access concepts of ordered 
 pairs whose external participants selected from their internal participants, as opposed to concepts 
 of ordered pairs whose internal participants selected from their external participants. But if the 
 internal argument of everyQ is a noun corresponding to a monadic concept of the usual sort—one 
 that can, at least in principle, apply to some but not all things—while the external argument of 
 everyQ is a polarized phrase that corresponds to a T-concept, that may create pressure in favor of 
 using the noun as the restrictor on external participants, as opposed to the restricted specifier of 
 external participants; see Pietroski (2005a, 2006) for related discussion in a framework that treats 
 sentences as expressions of type <t>.

Moreover, the index on a phrase like [everyQ guestN]Q2—corresponding to the (displaced) 
 aspect of a polarized phrase that is relevant to the abstraction associated with the external 
 argument of the quantifier—suggests an easy way of encoding the restriction that external 
 participants be among internal participants. We can restrict the permissible assignment variants. 
 Let’s say that for any assignments A and A*, and any index i, A*⊆A if: A* differs from A at 
 most in that A* does not assign to i everything that A assigns to i; whatever A assigns to i, A* 
 assigns one or more but perhaps not all of those things to i. Given an assignment A that assigns 
 (all and only) the guests to the first index, every restricted variant of A assigns one or more of 
 those guests to the first index. This puts number-neutrality to work in a perhaps unexpected way.

In section two, I posited the concept TARKSI{i, P} as a natural “semantic” concept whose 
 content can be specified formalistically with (41).
(41) \( \exists A^* : A^* \models \{ \text{ASSIGNS}(A^*, _, i) \} \) where \( \text{ASSIGNS}(A^*, _, i) \)

But perhaps the natural concept is \textbf{RESTRICTED Tarski} \{i, P\}. And perhaps (57)

(57) \( \exists A^* : A^* \subseteq A \{ \text{ASSIGNS}(A^*, _, i) \} \) & \text{SATISFIES}(A^*, P) \} \) where \( \text{ASSIGNS}(A^*, _, i) \)

is a better description of this concept's content. In any case, the instruction \textbf{ABSTRACT}(i, \mu)

can characterized in terms of \textbf{RESTRICTED Tarski} \{i, P\} instead of \textbf{Tarski} \{i, P\}. And given this

characterization, the proposed meaning specifications have the consequence that phrases like

[\text{every}_{Q \text{ guest}_N}] \text{Q2}

are restricted quantifiers, and that quantificational expressions exhibit the

“conservativity pattern” illustrated by the equivalence of (7) and (56).

(7) Every guest stabbed him.

(56) Every guest is a guest who stabbed him.

This still accommodates the meanings of (7-9) without generating an unwanted meaning for (10).

(8) She stabbed every guest.

(9) Most guests stabbed every spy.

(10) every guest who stabbed him

So by \textit{not} thinking of quantification in terms of Fregean typology, we can preserve the insights

of standard accounts while perhaps doing a little \textit{better} in terms of explaining the source of

\textit{constraints} on how human I-languages generate meanings.

4. Conclusion

A pattern emerges from the exercise conducted in this chapter. If one adheres to the idea that

combining expressions is fundamentally an instruction to construct \textit{conjunctive} concepts, along

with the idea that open class lexical items are instructions to fetch concepts with independent

content, one is led to say that certain aspects of syntax and various functional items are

instructions to \textit{convert} fetchable/constructable concepts into concepts that can be systematically

conjoined with others. Perhaps this is the \textit{raison d'être} of syntax that goes beyond mere recursive

concatenation: grammatical relations, like being the internal/external argument of a verb or
determiner, can carry a kind of significance that is intriguingly like the kind of significance that

prepositions have. These old ideas can be combined in a Minimalist setting devoted to asking

which conversion operations are required by a spare conception of the recursive composition

operations that human I-languages can invoke in directing concept assembly. The list of

operations surveyed here is surely both empirically inadequate, and yet already too rich. My aim

has been to offer a specific proposal as one illustration of minimalist thinking in semantics.
Chapter Seven: Summary

TBA
References (incomplete)


Austen, J. (1813). *Pride and Prejudice*.


Collins, J. (forthcoming). Three Grades of Syntactic Involvement
[chapter on Semantics and Sensible Ontology]


—(1987). Tacit knowledge and semantic theory:
Can a five per cent difference matter? Mind, 96, 441–462.


Erickson, K. (201x). UMD dissertation.


Linguistic Inquiry 1:429-38.
*Cognition* 28: 3-71.
*Pacific Philosophical Quarterly* 73: 31-43.
Graff-Fara (201x). You can call me stupid.


Hodges, W. — paper on compositionality.


Husserl, E.


Studies in the Linguistic Sciences 29.1.
(200x).
Kim, J. (1976). Events as Property Exemplifications. In M. Brandt & D. Walton (eds.)
(200x). paper on date names
Linguistics and Philosophy, 4:1-45
constitution. In Lexical Matters, CSLI Lecture Notes No. 24, I. Sag & A. Szabolcsi (eds).
Stanford, CA: CSLI Publications.
Cambridge, MA: Harvard University Press.
Lakoff, G. 1965: On The Nature of Syntactic Irregularity. Mathematical Linguistics and
University, Cambridge, Mass.
Cambridge, MA: MIT Press.
Semantics of Natural Language (Dordrecht: Reidel).
Philosophy of Science, Volume 7, Minneapolis: University of Minnesota Press, pp. 3–35.
Natural Language Semantics 19: 227-56.
Lidz, J. et.al. (201x). paper on verbal adicities and Gordon
Approach. In R. Bäuerle, et.al., eds., Meaning, Use, and Interpretation of Language
(Berlin: deGruyter).
—(2002). LF and natural logic. In G. Preyer and G. Peter (eds), Logical Form and
Donald Davidson on Truth, Meaning and the Mental (Oxford: Oxford University Press).
McCawley, J. (1968). Lexical insertion in a transformational grammar without deep structure.
Moltmann, F. (201x).
O’Shaughnessy, B. (1973). Trying (as the Mental ‘Pineal Gland’).
Journal of Philosophy 70: 365-86.
Pagin, P. — paper on compositionality
Synthese 21:320-34.
Linguistic Inquiry 6: 203-300.
—(197x).
Patterson, D. (201x). Book on Tarski,
Pelletier, J. — papers on compositionality
—papers on mass nouns, universal grinder
In Higginbotham, Pianesi, and Varzi 2000.
          Philosophical Issues 16: 221-45.
          15:157-190.
          M. Werning, W. Hinzen, & E. Machery (eds) Oxford Handbook of Compositionality.
          (Oxford: Oxford University Press).
          Verb Representation and Processing, R. de Almeida & C. Manouilidou (eds).

Pietroski, P. & Crain S. (2012). The Language Faculty. In The Handbook of Philosophy of
          Oxford: Oxford University Press.


          Twenty-First Century Psycholinguistics: Four Cornerstones, A. Cutler (ed.).
          Lawrence Erlbaum. Hillsdale, NJ.


          Oxford: Oxford University Press.

          Cambridge, MA: MIT Press.


Putnam, H. (197x). It Ain’t Necessarily So.
—(1975). The Meaning of ‘Meaning’. In K. Gunderson (ed.), Language,
          Mind and Knowledge (Minneapolis: University of Minnesota Press).

          (Oxford: Oxford University Press).

          Cambridge, MA: MIT Press.
— (201x). paper on intentional inexistent
Ricketts, T. (1996) Objectivity and Objecthood (?) Logic and Truth in Freges (?
— (192x). Logical Atomism
Saul, J.
—(200x). paper on scales.
Segal, G. (1991). In the mood for a semantic theory.
  *Proceedings of the Aristotelian Society* 91: 103–118.
— (1987). Direct Reference, Propositional Attitudes, and Semantic Content,
  *Philosophical Topics*, 15: 47–87.
Spelke, E. (2002). ‘Developing Knowledge of Space: Core Systems and New combinations’,
  in S. Kosslyn and A. Galaburda (eds.), Languages of the Brain (Cambridge,
  Mass.: Harvard University Press).
  Advances in the Investigation of Language and Thought (Cambridge, MA: MIT Press).
Stainton, R. (201x). ‘Meaning and Reference Some Chomskyan Themes’, in
  E. Lepore and B. Smith (eds.), Handbook of Philosophy of Language (Oxford:
  Oxford University Press).
— (201x). Historical essay with discussion of Montague/Kaplan
—Hannah/Esther paper…intro to collected essays?
Tarski, A. (1933). The Concept of Truth in Formalized Languages. Reprinted in *Logic,
  Indianapolis: Hackett.
  Research* 4: 341-75.
  J. Corcoran (Indianapolis: Hackett).
—(200x). *Occasion Sensitivity*.
Turing, A. (193x).
Uriagereka, J. et.al., eds., (200x). *Step By Step*.
—(199x). paper on central predicates in a theory of meaning
Zalta, E. essay on Frege’s Theorem in SEP.