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 uncontroversially.

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).[2]

(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 *un*ambiguous. 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.[3]

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 ***bɛr***. 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 ***bɛr***, 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 ***kaʊ***, 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 homophonous 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.[4] 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 easy 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 π 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 (198x) for relevant discussion. But I suspect that every 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 π has four or five meanings. But modulo vagueness, π will have *n* but not *n*+1 meanings, for some *n*; and in many cases where π is the pronunciation of a word-string, π 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.[5] 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'.[6]

(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.

---

[5] See Crain and Pietroski (2001, 2002) for discussion in the context of experimental evidence that young children already enforce adult-like constraints on homophony.

[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 thing 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 acquires 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 homophonies 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 (20α); and one way in which 'man with binoculars' is a constituent, as shown in (20β).

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

(20β) {[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 (20α) 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 (20α).

---

[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 (20α) 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 μ with a pronunciation π. 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".[8] 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 $E_1 \ldots E_n$ in a certain way W, then CE has a meaning of the *form* $\mu(W)[\mu(E_1)\ldots\mu(E_n)]$; where $\mu(\alpha)$ is the meaning of α, 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 E_1 and E_2 in a way that signifies a certain kind of conjunction, then $\mu(CE)$ can be the tripartite instruction **JOIN**$[\mu(E_1), \mu(E_2)]$; where this meaning is executed by applying a certain conjunctive operation to the results of executing the constituent meanings/instructions.[9] 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 E_1 and E_2 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 $\mu(E_1)$ and $\mu(E_2)$. As discussed in chapter two, $\mu(W)$ might be the operation of function-application, while $\mu(E_1)$ is a function from entities in a domain that includes $\mu(E_2)$ 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 (1892b) notion of *Sinn*; see also Church (1951), Horty (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 (20α) 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 (20a) as opposed to (20c).

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

(20b′) ∃x{Spy(x) & ∃y[Saw(x, y) & Man(y) & ∃z[Had(y, z) & Telescope(z)]]}

(20c′) #∃x{Spy(x) & ∃y[Saw(x, y) & Man(y)] & ∃z[Had(x, z) & 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.[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 '∃z[Had(x, z) & 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 '∃y[Saw(x, y) & 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 conjoining 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* had a telescope. English is not such a language. But if (27) is not like (28), what is it like?

(28) ∃y[Saw(x, y) & Man(y)] & ∃z[Had(x, z) & 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 (29b) rather than (29a)?

(29a) A spy with$_{\text{possessing}}$ a telescope saw a man.

(29b) A spy with$_{\text{by-using}}$ a telescope saw a man.

---

[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(SE_1), μ(SE_2)]' by rewriting with dyadic/restrictable quantifiers:

∃x[Spy(x), ∃y[&[Man(y), ∃z[Telescope(z), Had(y, z)]], Saw(x, y)]];

∃x[Spy(x), &[∃y[Man(y), Saw(x, y)], ∃z[Telescope(z), 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 μ([[saw [a man]] [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 μ([[saw [a man]] [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 distracts. 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 (20α) means.

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

If the goal is to specify the meaning of (20α) in a way that shows how its meaning is—or at least could be—determined by its lexical items and grammatical structure, then it helps to think of 'with a telescope' as a predicate of events, as suggested by the regimentation (20aa′);

(20aa′) ∃e∃x{Spy(x) & ∃y[See(e, x, y) & Past(e) & Man(y) &
∃z[With(e, z) & Telescope(z)]]}

where 'See(e, x, y)' is true of ordered triples <ε, α, β> such that ε is an event of α seeing β. From this perspective, 'saw a man' is unlike '∃y[Saw(x, y) & Man(y)]' and more like '∃y[See(e, x, y) & Past(e) & Man(y)]', which is true of ordered pairs <ε, α> such that ε is an event of α seeing a man. The meaning of (20β) can be specified in similar fashion.

(20bb′) ∃e∃x{Spy(x) & ∃y[See(e, x, y) & Past(e) & Man(y) &
∃z[With(y, z) & Telescope(z)]]}

That is a step in a good direction. As discussed below, it also lets us start to explain the sense in which (20α) 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 '∃y[See(e, x, y) & Past(e) & Man(y) & ∃z[With(x, z) & 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′) ∃e∃x{Spy(x) & ∃y[See(e, x, y) & Past(e) & Man(y) &
∃z[With(x, z) & Telescope(z)]]}

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

[11] *Pace* Horwich (1997, 1998), it's not enough to display (20α) while saying that its elements and modes of combination mean what they do; see Pietroski (2000b, 2005b) for extended discussion.

(20aa″) ∃e{∃x[Experiencer(e, x) & Spy(x)] &
∃y[Psee(e, y) & Past(e) & Man(y) & ∃z[With(e, z) & Telescope(z)]]}

where 'Psee(e, y)', 'P' connoting 'Passive', is true of pairs <ε, β> such that ε is an event of seeing β. On this view, 'saw a man' is more like '∃y[Psee(e, y) & Past(e) & 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 'Psee(e, y)' with 'Psee(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 'Psee(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 'Psee(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 μ('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 μ('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 pre-theoretic 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 μ with a pronunciation π; 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 H_2O 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 H_2O (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 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 β and γ 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 (*Sinn*) and semantic value (*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, $\{ \ldots, (-2, 3), (-1, 2), (0, 1), (1, 0), (2, 1), \ldots\}$, 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 λ-δ-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 λ-*K*-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 λ-δ-conversion, where the notion of identity of functions is introduced into the system by the symbol δ, *it is necessary, in order to preserve the finitary character of the transformation rules,* so to formulate these rules that an *interpretation by functions in extension becomes impossible*. The expressions which appear in the calculus of λ-δ-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.[14] 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.[15]

---

[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) $\exists x[Fx \ \& \ \forall y[Rxy \supset \exists z[Szx]]$

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 $<\mu, \pi, n>$; where n is a number, $0 < n \leq 1$, 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 $<\mu, \pi, n>$ 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 <μ, π> 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 "reversible" 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 if 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; cp. 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.[19]

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

[19] 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 (ReflOnLg).[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.[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 `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?

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 Γ to be used by a population 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 Γ is used by a population P, whereas another grammar Γ', which generates the same language as Γ, is not. I have tried to say how there are facts about P which objectively select the languages used by P. I am not sure there are facts about P which objectively select privileged grammars for those languages...a convention of truthfulness and trust in Γ will also be a convention of truthfulness and trust in Γ' whenever Γ and Γ' generate the same language (p. 20).

[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.[22] 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 communciation," 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 ***kaʊ*** 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/arbitrary 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 **saɪprəs** 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 **saɪprəs** to refer to cypresses, perhaps allowing for "deferential" concepts of whatever certain experts think about when using a concept they have linked to **saɪprəs**; 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) Sentence ➔ Subject copula Predicate
Subject ➔ Quantifier Predicate
Subject ➔ Name
Predicate ➔ 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 ➔ 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 Γ is used by P if and only if Γ is a best grammar for a language £ that is used by P in virtue of a convention in P of "truth and trustfulness in £.[23] One might have thought that a "best" grammar for the alleged set £ 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."

[23] 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 hypostatizing 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/I-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 explication.

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*) ≡ Ψ(*c*)'; where '*c*' ranges over contexts of some kind, and '≡' 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) ≡ p'; where 'p' can be the form 'for each sequence σ of potential values for any variables, Ψ(σ)'. 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*) ≡ 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 'μ(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.[26] 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-evaluable 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 Fregean 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 (195x). 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 comprehensibly—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 I-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 gleelessly 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 justify 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.
<u>Venice was a better location.</u>
Torcello was moved to a better location.

(64+) Torcello was moved to Venice.
Venice was a better location.
<u>Venice may need to be moved.</u>
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 a 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'. Analagous 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) True('Bert yells.') ≡ Yells(Bert)
(77-T) True('Ernie snores.') ≡ Snores(Ernie)

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

(71-T*) True('Bert yells.') ≡ Snores(Ernie)
(77-T*) True('Ernie snores.') ≡ 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) Yells(Bert) ≡ 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) True('Bert yells.') ≡ Yells(Bert) & ⊢

where ⊢ 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**
can be 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) for each context $\mathbf{c}$, True(S, $\mathbf{c}$) ≡ F($\mathbf{c}$)

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) form 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) 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 *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)

(83) 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 I-language

advocates of (D) need to provide reasons for thinking that (B) is true.

(B) each human language **H** is a biologically implementable I-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) I-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 unambiguous yet 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 (196x) 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, iif 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 a 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 explanadum 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 explananda. 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 "disquotational" 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 x 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.