Paul M. Pietroski

Meanings via Syntactic Structures

Chomsky (1957) offered prescient suggestions about how to formulate theories of understanding for the spoken languages that human children can naturally acquire. We can view his proposal as a prolegomenon to a theory of meaning that combines a layered theory of syntax with an account of how humans can naturally use expressions in acts of referring, asserting, querying, and so on; cp. Austin (1961); Davidson (1967); Dummett (1976); Higginbotham (1985). Though crucially, Chomsky's (1957) conception of reference differed from more familiar conceptions of denotation, as in Frege (1892b) or Russell (1905, 1912, 1957); cp. Goodman (1949, 1953), Strawson (1950), Chomsky (1977, 1995b, 2000b, 2017). I think we should develop the kind of semantic theory towards which Chomsky has long been pointing; see Pietroski (2017abc). But in this essay, my aim is simply to articulate a sense in which *Syntactic Structures* outlined a program for semantics, well before Davidson (1967), Montague (1970), and Lewis (1970).

1 Introduction

Human children come equipped with a remarkable resource. Given an ordinary course of experience, an ordinary child can acquire at least one procedure that generates unboundedly many linguistic expressions that are pronounceable and meaningful. Children also tend to acquire procedures that are remarkably similar to those acquired by local peers, even though each child's experience is limited and somewhat idiosyncratic. Correlatively, adults in a linguistic community will agree—if asked—about many subtle linguistic facts.

For example, mature speakers of English can tell that (1) is ambiguous,

(1) the old man was frightened by the new methods

even if they have not previously encountered this string of words. We can hear (1) as a sentence according to which the old man found it scarily hard to deal with some new ways of doing something (e.g., getting books from the library). But on another construal, perhaps less obvious, (1) is a sentence according to which new

methods of frightening people were successfully used on the old man.[1] Interestingly, (2) can be understood as a description of a certain man,

(2) the old man frightened by the new methods

though not as a sentence; while (3) can only be understood as a sentence with a bizarre meaning,

(3) the old man frightened the new methods

and (4) is a mere string of words with no unified meaning.

(4) old the was man by frightened new the methods

A monolingual speaker of Japanese will not know such facts, unless she is explicitly told about them—say, in a linguistics class conducted in Japanese. But such a speaker will also fail to know, with regard to (1), what the words and any constituent phrases mean. The important point here is that a speaker of English who knows what the words mean, and how they can be fit together, can recognize that (1) is ambiguous and (4) is gibberish.

A caveat: using words like 'English' and 'Japanese' in this way reflects an idealization that is not always appropriate. The many competent speakers of English do not share a common procedure that connects a public stock of meanings with the very same pronunciations. But in many contexts, including this one, we can abstract away from the relatively minor variations exhibited by the English procedures acquired in London, Sydney, Glasgow, Chicago, etc. Then we can pretend that English is a single language (with a fixed vocabulary). Likewise, we might treat the sundry Japanese procedures as equivalent, perhaps to highlight more dramatic contrasts between any one of them and *any* English procedure. For despite these contrasts, a human infant is prepared to acquire English or Japanese or both, without any special instruction regarding facts like those illustrated with (1-4). So even if we speak of variation across dialects of *a* language — maintaining the pretense that English is *one*, and Japanese another — acquiring a language like English or Japanese is a matter of acquiring a generative procedure of a special kind.

1 The meanings differ, even if the new way of getting books was designed to frighten the elderly. I'll return to this example of what Chomsky (1957, pp. 86-90) calls *constructional homonymity*; see his example (114).

Chomsky (1986) sharpened this point by speaking of I-languages, which he defined as generative procedures, in explicit contrast with sets of expressions. Given this technical notion, one can ask how the I-languages that children naturally acquire *differ* from other procedures that connect signals of some kind with interpretations of some kind. For reasons noted below, Chomsky (1957, p.13) said that a language is a set of sentences, each of which is "finite in length and constructed out of a finite set of elements." But as the restriction and subsequent discussion makes clear, the real issues concerned methods of generation, not sets of generable sentences. More specifically, the assumption was that among the computational procedures characterized by Turing (1936) and Church (1941), there are some that humans can acquire and use to connect pronunciations with meanings—whatever pronunciations and meanings are—in ways that children regard as natural, at least for purposes of formulating and evaluating the relevant options in the course of acquiring languages like English or Japanese.[2]

This invites a research project that Chomsky (1957) initiated. Given a vocabulary that lets us describe the full range of computable procedures, we can try to provide a rough description of the smaller range that corresponds to the spoken languages that humans can naturally acquire. But the terminology we initially use, as we start to delimit the relevant subclass of procedures, is likely to be artificial in the following sense: this subclass may well be arbitrary, at least to some degree, from an abstract computational perspective that ignores requirements imposed by the need for biological implementations in human children; and while part of our initial aim (as theorists) is to locate certain procedures in a larger space that includes many logically coherent options that are *un*available to children, the targeted procedures form a scientifically interesting subclass because they are the ones available to children. So if all goes well, we should expect our theoretical vocabulary to evolve — from Subject/Copula/Predicate, to NP/Aux/VP, to whatever— as linguists make progress towards specifying certain naturally acquirable procedures in terms of notions that are naturally available to children.[3]

2 We can speak of procedures as abstract mappings from elements of a domain to elements of a range. But at least when the functions in question are Turing-computable, we can also speak of physical implementations of the mappings; and implementations of a procedure/program—e.g., as a string of binary numerals, suitably stored on a certain machine—can be acquired and used in various ways. Chomsky's (1986) I-language/E-language contrast echoed Church's (1941) contrast between functions-in-intension and functions-in-extension; cp. Frege (1892a).

3 Of course, one can hypothesize that the computational character of these procedures is not entirely arbitrary; see Chomsky (1995a, 2000a). And in principle, lucky inquirers might quickly hit on the right terminology for their subject matter; though in actual scientific practice, this never happens. So given that semanticists often use Church's (1941) lambda calculus to describe mean-

Correspondingly, the goal is not to specify and study certain sets of sentences. The goal is to describe the cognitive resources that humans deploy in the acquisition and use of certain generative procedures. Though in terms of making progress, we face an immediate problem, absent clear ideas about what meanings are.

It's hard enough to say what pronunciations are, even if we focus on spoken as opposed to signed languages. But at least we can talk about sequences of phonemes. If we want to talk about *expressions*, however, the phenomenon of homonymy introduces a familiar complication. Consider (5-8); where only (5) and (6) are strings that correspond to meaningful phrases.

(5) by the new methods

(6) buy the new methods

(7) by the knew methods

(8) buy the knew methods

In many cases, a sequence of phonemes corresponds to more than one word. Still, one might hope to confine this kind of homonymity to a stock of finitely many atomic expressions that can be distinguished formally — i.e., without appeal to their meanings — as when lexicographers use superscripts to distinguish the several words that share the pronunciation indicated with 'bank' or /baeŋk/. But even given a specification of the meaningful sequences of English phonemes, along with a way of disambiguating lexical homonyms, how can we start to describe the procedures that connect these sequences with their meanings if we don't know what meanings are?[4]

---

ings, we should remember that this formalism was invented as a tool for representing the entire space of computable functions. Likewise, the now familiar Fregean hierarchy of types—rooted in basic types that correspond to denotable entities and a pair of truth values—was introduced in the course of inventing a language that can be used to encode mathematical thoughts that humans can entertain, with help, even if the thoughts cannot be expressed in natural languages; see Pietroski (2017b) for discussion.

4 In my view, semanticists still don't have good analogs of phoneme sequences. Invoking sets of possible worlds is obviously inadequate, absent an independent characterization of worlds *and* a method for pairing distinct sets of worlds with nonsynonymous sentences—e.g., 'If there are groundhogs on Venus, then five is an even number' and 'There are finitely many prime numbers if woodchucks exist on Hesperus'.

2 Ambiguities and Analyses

Chomsky's (1957) strategy was inspired. But we should remember how bold it was. As an initial model that he soon modified, Chomsky *replaced* the question of how a computational device could generate (all and only) the pronunciation-meaning pairs of English with the somewhat clearer question of how a computational device could generate certain strings of English words—the "grammatical" strings—without also generating boundlessly many others. This was helpful, in part because given the new question, it was easy to show that some simple devices would overgenerate and/or undergenerate. But the grammatical strings served as first-pass proxies for the corresponding sentential meanings; and as discussed below, equivalence classes of derivational histories—represented as trees—provided better second-pass proxies.

Chomsky knew that a single string of words can correspond to more than one sentence meaning, much as a single string of phonemes can correspond to more than one word meaning. So he introduced a general notion of "constructional homonymity" for cases in which a phoneme sequence "is analyzed in more than one way on some level" (p. 86).

> For example, for many English speakers the phoneme sequence /aneym/ can be understood ambiguously as either 'a name' or 'an aim'. If our grammar were a one-level system dealing only with phonemes, we would have no explanation for this fact. But when we develop the level of morphological representation, we find that, for quite independent reasons, we are forced to set up morphemes 'a', 'an', 'aim' and 'name', associated with the phonemic shapes /a/, /an/, /eym/ and /neym/. Hence, as an automatic consequence of the attempt to set up the morphology in the simplest possible way we find that the phoneme sequence /aneym/ is ambiguously represented on the morphological level (p. 85).

Similarly, even if this ambiguity is resolved in favor of 'a name', intuition suggests that string (9) is ambiguous in a now familiar way indicated with (9a) and (9b);

(9) referred to a star with a name
  a. [$_{VP}$ referred to a [$_{NP}$ star with a name]]
  b. [$_{VP}$ [referred to a star] [$_{PP}$ with a name]]

where 'star with a name' is a constituent on analysis (a), and 'referred to a star' is a constituent on analysis (b). Compare 'saw a teacher with a telescope', which can be understood as implying that some teacher *had* a telescope, or as implying that someone *used* a telescope to see a teacher.

Example (1) is subtler, in part because 'frightened' is morphologically complex.

(1) the old man was frightened by the new methods

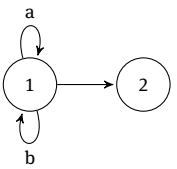
But the two meanings seem to reflect distinct ways of combining the meaningful elements. Given the details in Chomsky (1957), including his account of passive constructions in terms of transformations (pp. 42-43), one might describe the ambiguity of (1) by positing two Deep Structures that correspond to distinct thoughts about the old man: the new methods (were things that) frightened him; or the new methods were used (by someone) to frighten him. Subsequent work led to other proposals; see, e.g., Pesetsky (1995). But here, the details are less important than the methodological moral Chomsky drew from the phenomenon of nonlexical ambiguity.

> This suggests a criterion of adequacy for grammars. We can test the adequacy of a given grammar by asking whether or not each case of constructional homonymity is a real case of ambiguity and each case of the proper kind of ambiguity is actually a case of constructional homonymity (p.86).

This passage highlights an important point, especially in light of the surrounding chapter ("The Explanatory Power of Linguistic Theory," discussed by Lidz in this volume). In providing a partial grammar for English, Chomsky's goal was not to partially specify a procedure that generates all and only the phoneme sequences that can be understood as English sentences. The aim was to characterize a procedure that generates the English pronunciation-meaning pairs *without also* generating many boundlessly other such pairs; and this was in the service of the larger goal of characterizing such a procedure as one possible choice from the range of diverse but fundamentally similar linguistic options that are available to human children.

The initial presentation in terms of word strings was a pedagogical device—a way of introducing students to the subject. Given the words of English, it's easy to imagine a finite state automaton that yields all the corresponding word strings, and hence each one that can be understood as an English sentence.[5]

5 The procedure below generates every string that can be formed from a lexicon consisting of 'a' and 'b'.



With regard to combining meanings, we (still) haven't discovered the relevant analog of concatenating inscriptions. But we can imagine supplementing a universal string generator with a "semantic procedure" that connects each string of atomic expressions with all the sentence meanings that could be formed from those atoms. Then however one describes the grammatical strings, each one would be connected to the corresponding sentence meanings, though at the cost of massive overgeneration.

Chomsky (1964, 1965) stresses these points more explicitly, using contrastive examples like (10) and (11); where '#' indicates a coherent but unavailable construal of the string.

(10) John is easy to please
  a. It is easy for us to please John.
  b. # It is easy for John to please us.

(11) John is eager to please
  a. # John is eager for us to please him.
  b. John is eager that he please us.

Even more interestingly, (12) is three ways ambiguous;

(12) the boy saw the woman walking towards the church
  a. The boy saw the woman while (he was) walking towards the church.
  b. The boy saw the woman who was walking towards the church.
  c. The boy saw the woman walk towards the church.

and unlike (12b), (12c) implies that the boy *saw an event* of the woman walking. But (13) is unambiguous, having only the interpretation that corresponds to (12c).

(13) this is the church the boy saw the woman walking towards

We understand (13) as meaning that the church is such that the boy saw the woman *walk* towards it. Yet superficially, (13) is more like (12) and the equally ambiguous (14).[6]

6 As Chomsky (1964) also notes, 'what did Mary see the boy walking towards' can be used to ask which thing is such that Mary saw the boy *walk* towards it—but not which thing is such that Mary saw the boy *who was walking* towards it, or such that Mary saw the boy *while walking* towards it. This suggests a common constraint on the interpretation of relative clauses and questions; see Ross (1967). For discussion of the relevance of event variables in accounting for constraints on

(14) this is the church such that the boy saw the woman walking towards it

Similar remarks apply to (15), one of Chomsky's (1957) examples of constructional homophony.

(15) John found the boy studying in the library

We could invent languages that generate (10-15) in ways that are subject to different constraints—e.g., languages in which (11) is as ambiguous as 'the duck is ready to eat', and (10) has only its English meaning, or vice versa. But naturally acquirable languages connect pronunciations with meanings as they do. Even controlling for morphological and lexical homonymy, a string of (formally disambiguated) words can still be constructionally homonymous in constrained ways. In many cases, it may be hard to know how many ways a given string can be understood. But prima facie, the facts to be accounted for include boundlessly many instances of the following generalization: each grammatical string of English words is constructionally homonymous in exactly n ways for some n, $n \geq 0$. While (12) is three ways ambiguous, it cannot be understood in the fourth way indicated with (12d).

(12) d. # The boy saw the woman and walked towards the church.

So given Chomsky's proposed criterion of adequacy, a grammar for English cannot merely specify the grammatical strings of phonemes. Each string must be paired with the right number of analyses; and in the case of ambiguous strings, it needs to be plausible that each way of understanding the string is reflected by one of the analyses. But the point is not confined to grammatical strings. For example, (16) is degraded but comprehensible;

(16) the child seems sleeping

cp. (17), which is bizarre yet grammatical.

(17) colorless green ideas sleep furiously

We understand (16) as meaning that the child seems to be sleeping, and not as meaning that the child seems sleepy; see also Higginbotham (1983). So we want a grammar of English to yield a suitable analysis of (16).

---

homonymity, and arguments that the grammatical strings do not form an interesting class, see Higginbotham (1985) and Pietroski (2017c).

Correlatively, the mere ungrammaticality of (18)

(18) * she have might been there

does not yet explain why (18) cannot be understood as synonymous with (19).

(19) she might have been there

The explanandum is not merely that (18) is *defective*, but rather, that it cannot be understood as a sentence—even though the words in (18) can be combined to form a perfectly fine sentence. A good grammar for English should assign some analysis to (19) *and not* assign an equivalent analysis to (18). Seen in this light, examples of constructional homonymity are special cases of strings that have more than zero meanings, and hence strings that should be assigned more than zero grammatical analyses.[7]

Chomsky (1957) left room for levels of semantic analysis other than Deep Structure and Surface Structure. He also noted that (20) and (21), on their obvious construals,

(20) everyone in the room knows at least two languages

(21) two languages are known by everyone in the room

are not remotely equivalent. The former, but not the latter, can be used to describe a situation in which no two people in the room know the same language. Such examples suggest the possibility of "covert" transformations of Surface Structures; see, e.g., May (1985). But given endlessly many examples like (22) and (23),

(22) she could haven't being fly
  a. # She could not have been flying.

(23) Romeo loves Juliet
  a. # Juliet is such that Romeo is such that she loves him.

theorists must be careful in adding a new level of analysis. The more complex grammar must not generate unattested pronunciation-meaning pairs that are not generated by the simpler grammar.

7 See Berwick et al. (2011) for discussion in the context of responses to corresponding "poverty of stimulus" arguments.

A related point is that if one allows for more than one meaning per grammatical analysis, then absent an independent proposal about how analyses are (not) related to meanings, the whole project of providing and evaluating constrained grammars is threatened. If a single analysis can correspond to two or more meanings, then providing a grammar that assigns *n* analyses to a string with *n* meanings does yet explain why the string fails to be more ambiguous than it is.

From this perspective, one shouldn't exclude the possibility of a semantic theory that pairs Surface Structures with hypothesized mental representations, as opposed to formal specifications of truth conditions; see Katz and Postal (1964); Katz and Fodor (1963). If theories of meaning are theories of understanding, then *pace* Lewis (1970), one cannot mandate that such theories be truth-theoretic; see Harman (1974). But if the mental representations are said to be the meanings that grammars connect with pronunciations, then the mapping from grammatical analyses to the posited meanings must be suitably constrained to avoid overgeneration. For example, an algorithm that pairs Surface Structures with Katzian disambiguaters must not yield too many meanings for (10-23). And if one posits a level of *grammatical* structure that is specified in terms of meanings, then analyses will not disambiguate strings formally, raising the question of how analyses are individuated for purposes of deciding whether or not a theory assigns *n* analyses to a given string. Put another way, theorists can posit levels of grammatical analysis that reflect meaning. But if analyses are to serve as proxies for meanings, the former can't be individuated in terms of the latter.[8]

8 Montague (1970) adopted a very different view without discussing the issues regarding overgeneration.

> Like Donald Davidson I regard 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; and the developments emanating from the Massachusetts Institute of Technology offer little promise towards that end (p. 188).

The talk of "truth under an arbitrary interpretation" suggests misunderstanding of Davidson's (1967) project; see Lepore (1983). More importantly, Montague's conception of serious syntax was remarkably implausible, and yet the basis for his influential decision to ignore Chomsky's (1957; 1964; 1965) conception of how a formal theory of syntax could be part of a substantive theory of understanding. It might be useful to ask how the field would have developed if Montague had been set aside, as irrelevant to the study of natural language, given the arguments (based on actual phenomena) that had emanated from MIT. Cp. Davidson (1967, p. 315), "Recent work by Chomsky and others is doing much to bring the complexities of natural languages within the scope of serious semantic theory." While Davidson's conception of reference differed from Chomsky's, this could have led to fruitful discussion, with less emphasis on mapping word-strings to expressions of the lambda calculus; cp. Lewis (1970, 1975).

3 Trees and Transformations

Following Chomsky's lead, I began with some mundane observations about children, and then noted some examples of homonymy. Instead of *starting* with assumptions about what meanings are—and how they are related to use, truth, logic, reference, functions, conventions, possible worlds, or other things—one can initially ground theoretical talk of meaning in specific facts about how many meanings certain strings of phonemes have, and more generally, how natural languages like English do and don't connect pronunciations with meanings. This highlights empirical phenomena that turned out to be theoretically tractable. But given this "syntax first" approach to semantics, it's important to be clear about what expressions and analyses are. In particular, if the idea is to use analyses as formal proxies for meanings—thereby delaying vexed questions about what meanings are, and how to connect a pronunciation with exactly *one* of them—we need a way of deciding how many analyses a proposed grammar assigns to a given sequence of phonemes. Otherwise, we can't assess the degree to which a proposed grammar is (in)adequate by the proposed standards.

If only to fix terminology in a simple way for his initial discussion, Chomsky (1957) took sentences to be strings as opposed to structured objects. Instead of saying that (1)

(1) the old man was frightened by the new methods

is homonymous as between two different sentences, he regarded string (1) as a sentence that can be classified in two different ways, each of which corresponds to a certain diagram (or "tree") that represents a certain partial ordering of the relevant lexical items. But the grammar that Chomsky provided did not yield such diagrams, or any other structured objects, as *outputs*. The derivational procedure he described yielded certain strings, and thus distinguished those strings from underivable alternatives.[9] Though typically, there is more than one legitimate derivation of a string, even ignoring homonymy.

As an illustration, consider (23) and the toy grammar (24).

(23) Romeo loves Juliet

9 The tacit assumption was that being derivable (from a naturally acquirable grammar) is a formally decidable property of strings; cp. being a tautology of first-order logic.

(24) S → NP VP
NP → Romeo
NP → Juliet
VP → V NP
V → V loves

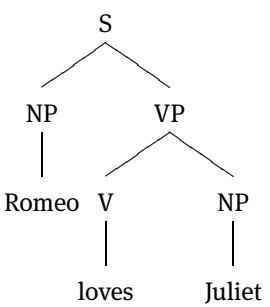
The grammar licenses several derivations of (23), including the two shown below.

(25)
S
NP VP
Romeo VP
Romeo V NP
Romeo loves NP
Romeo loves Juliet

(26)
S
NP VP
NP V NP
NP V Juliet
NP loves Juliet
Romeo loves Juliet

But (25) and (26) correspond to the same "collapsed PS tree," depicted as (27).

(27)



This tree preserves key information about which rules were applied how many times, while abstracting from differences of order across derivations; see Lasnik (1999) for lucid discussion. Note that (25-27) all capture the idea that relative to grammar (24), string (23) *is an* S in which 'loves Juliet' *is a* VP.

On a first reading, one might wonder why Chomsky fussed about defining a particular notion of derivational equivalence. But one important reason was to be explicit about the proxies for meanings, which could not be plausibly identified with strings or derivations. The idea was that (15) has two *analyses*, not that it has two derivations.

(15) John found the boy studying in the library

And while (23) has more than one derivation, it doesn't have more than one meaning.

Focusing on equivalence classes of derivations (cp. Goodman (1951)) was a beautiful idea that offered a new way of thinking about how meanings are related to technical notions of extensional equivalence. It was also an empirically fruitful idea about what matters for grammatical distinctions; see Lasnik and Kupin (1977) for discussion and an independent specification of derivational equivalence. But in terms of providing proxies for meanings, the key point is that given a way of saying which derivations are equivalent, Chomsky could identify analyses with classes of equivalent derivations and then say that cases of constructional homonymity are cases in which a string of phonemes has *non*equivalent derivations.

On this view, a formal theory of syntax for a language L—a theory that specifies, without appeal to meanings, how expressions of L are generated—can be a core component of a broader theory of how speakers of L understand expressions of L. Moreover, a theory of syntax for L can help explain how speakers understand expressions of L in structured ways, even if the proposed grammar does not itself generate structured objects.

In distinguishing phrase structure grammars from finite state systems, Chomsky's point was not that the former let us build certain tree-like objects. His aim was to focus on procedures that can at least generate strings in which distant elements were introduced together, and then motivate appeals to transformations (of some kind) in specifying how humans naturally generate linguistic expressions, while recognizing that invoking transformations can easily lead to overgeneration. As Chomsky (1957, p.23) puts it, "there are processes of sentence-formation that finite state grammars are intrinsically not equipped to handle." But likewise, we need to posit grammars that generate strings in ways that phrase structure grammars do not. So we face the empirical question of which other procedures should be invoked for which constructions. And when evaluating proposals, considerations of simplicity are relevant.

> ...if we have a transformation that simplifies the grammar and leads from sentences to sentences in a large number of cases (i.e., a transformation under which the set of grammatical sentences is very nearly closed), then we attempt to assign constituent structure to sentences in such a way that this transformation always leads to grammatical sentences, thus simplifying the grammar even further (p. 83).

Space constraints preclude discussion of the many implications of this point for semantics. But let me offer a reminder that Chomsky's analysis of English auxiliaries, as in (28),

(28) she could have been flying

addressed ancient questions concerning how modes of speech are related to thoughts that seem to exhibit Subject-Copula-Predicate structure. We can describe (28) as a result of transforming a string that exhibits NP-Aux-VP structure, as shown in (28D), into something more like (28S).

(28D) $[_{S}[_{NP}$ she] $[_{Aux}$ *past* can have-en-be-ing] $[_{VP}$ fly]]

(28S) $[_{S}$ $[_{NP}$ she] $[_{Aux}$ can-*past* have be-en] $[_{VP}$ fly-ing]]

So we can describe (28) as a generable string that corresponds, via slight grammatical detour, to a thought whose complex Copula links a Subject to a Predicate. Similar remarks apply to (29),

(29) she flew
D $[_{S}$ $[_{NP}$ she] $[_{Aux}$ *past*] $[_{VP}$ fly]]
S $[_{S}$ $[_{NP}$ she] $[_{Aux}$] $[_{VP}$ fly-*past*]]

and to (30), given Chomsky's (1957) proposal about 'do' supporting the tense morpheme.

(30) she did fly
D $[_{S}$ $[_{NP}$ she] $[_{Aux}$ *past*] $[_{VP}$ fly]]
S $[_{S}$ $[_{NP}$ she] $[_{Aux}$ do-*past*] $[_{VP}$ fly]]

As Chomsky notes, further computation is required to convert Surface Structures into phonemic sequences. More importantly, and as one would expect from the quote above, he treats negation as a transformation: a rule, in this case optional, that converts (29D) into a negated form—she *past*+n't fly—which is converted to (31) via "Do-Support" and morphology (p. 62).

(31) she didn't fly

In this way, we can also describe (31) and (32)

(32) she couldn't have been flying

as generable strings that correspond to Subject-Copula-Predicate thoughts.[10] This is independently interesting because natural language negation doesn't seem to be a prefix that attaches to complete sentences, much less an expression of the Fregean type <**t**, **t**>; see Horn (1989) for extended discussion. We can't use (33) to express the thought we can voice with (34).

(33) not she flew

(34) it isn't the case that she flew

Morphemes that signify negation seem to be interwoven with tense and modal auxiliaries, as Aristotle suggested; though as he observed, this leads to puzzles. For example, (35)

(35) Socrates may not speak

can be used to say that it is impermissible for Socrates to speak. We still don't know why.

4 Rules and Reference

Thinking of negation as an optional transformation invites appeal to Chomsky's (1957, p. 45) distinction between the *kernel* of a language, as specified by some grammar G—viz., the set of sentences generated by applying obligatory transformations to the terminal strings of the phrase-structure component of G—and the other sentences generated by G. And if grammars support this distinction, one shouldn't assume that theories of meaning ignore it.

In specifying how strings are understood relative to a grammatical analysis, we may well posit lexical meanings and combinatorial operations that categori-

10 In a different idiom, one might describe *n't* as a functional element that is attracted to tense—see Laka (1994)—with the result that in (31), *n't* blocks tense from hopping further to the right, thus triggering Do-Support; cp. (29).

cally determine many phrasal meanings that correspond to nodes of collapsed PS-trees. But we may also be led to posit various syncategorematic operations—akin to Tarski's (1933; 1944) treatments of negation, connectives, and quantifiers—that convert one meaning into another, in accord with a rule, but not via composition in any interesting sense.[11] In particular, one can imagine rules according to which each instance of NP-Aux-VP has a meaning that is determined by (i) the meanings of the constituent NP and VP, and (ii) the particular connector, which may comprise two or more auxiliary elements; where (ii) need not have a meaning that is determined, node by node, relative to the corresponding equivalence class of derivations. Other things equal, one might want grammars that "minimize mismatches" among meanings, derivations, and pronunciations; see Barker and Jacobson (2007). But other desiderata bear on the questions Chomsky highlighted.

Let me stress that describing expressions as strings (relative to analyses) leaves room for the idea that meanings and/or pronunciations are structured things, perhaps generated via simple operations that don't involve transformations; see Idsardi (forthcoming), Pietroski (forthcoming). In particular, a syntactician can treat equivalence classes of derivations as proxies for meanings, while thinking that psychologically realized grammars connect phonemic strings with structured meanings of some kind. But it is far from the spirit of Chomsky (1957) to suppose that grammars generate labeled phrase markers that await interpretation, like the well-formed formulae of an invented formal language. The idea was not that grammars generate uninterpreted syntactic structures, but rather, that humans enjoy a grammatical competence that lets us connect meanings with pronunciations in (autonomous) ways.[12]

Correlatively, as Chomsky (1957, pp. 103-4) emphasized,

11 Cp. Chomsky's discussion of 'and' on pp. 35-36, and his similar formulation of a rule concerning 'Aux' on p. 43.

12 In describing his efforts to characterize the syntax of English, Chomsky spoke of "studying language as an instrument or a tool, attempting to describe its structure with no explicit reference to the way in which this instrument is put to use." His motivation "for this self-imposed formality requirement for grammars" is that "there seems to be no other basis that will yield a rigorous, effective, and revealing theory of linguistic structure." Though as he immediately stressed, the formality requirement is "perfectly compatible with the desire to formulate it in such a way as to have suggestive and significant interconnections with a parallel semantic theory (p.103)." Moreover, even if one rejects Chomsky's empirically motivated reasons for his formality requirement, violating it raises questions about how else one can distinguish psychologically realized grammars from all the other aspects of human cognition that influence the use of linguistic expressions.

> To understand a sentence we must know much more than the analysis of this sentence on each linguistic level. We must also know the reference and meaning of the morphemes or words of which it is composed; naturally, grammar cannot be expected to be of much help here. These notions form the subject matter for semantics.

And in a footnote attached to the word 'meaning', Chomsky cites Goodman (1949, 1953).

> Goodman has argued—to my mind, quite convincingly—that the notion of meaning of words can at least in part be reduced to that of reference of expressions containing these words....Goodman's approach amounts to reformulating a part of the theory of meaning in the much clearer terms of the theory of reference, just as much of our discussion can be understood as suggesting a reformulation of parts of the theory of meaning that deal with so-called "structural meaning" in terms of the completely nonsemantic theory of grammatical structure. Part of the difficulty with the theory of meaning is that "meaning" tends to be used as a catch-all term to include every aspect of language that we know very little about. Insofar as this is correct, we can expect various aspects of this theory to be claimed by other approaches to language in the course of their development.

Chomsky assumed that his professional readers would know about Goodman's work, or follow the citations to it, and recognize that the intended notion of reference was not Russell's (1912) notion of denotation or Frege's (1892b) notion of *Bedeutung*. But given that appeals to denotational "semantic values" have become standard, it may be useful to recall the context—Cambridge Mass. in the 1950s—along with Goodman's main point and his central example.

Goodman assumed that (a) linguistic expressions have interpretations of some kind, but (b) our ordinary notions of meaning are not themselves suitable for use in scientific theories, and (c) simply identifying interpretations/meanings with denotations/extensions is not the way forward. The suggestion was to develop a notion of reference condition such that expressions with the same reference condition—i.e., expressions that impose the same conditions on acts of reference (or predication) across constructions—could be said to have the same meaning, at least to a first approximation that could serve tolerably well in theories of linguistic understanding. Today, the project of "reconstructing" talk of meanings might seem quaint. But fads change.

Prior to Tarski (1933, 1944), one might have thought that *any* substantive notion of "interpretation" will be too vague and interest-relative to do duty in a scientific theory. Given Tarski's discussion of truth, however, matters became quite unclear with regard to meaning. On the one hand, there was pressure to eschew notions of meaning in favor of extensional notions, which apply to whatever things they apply to regardless of how those things are described or depicted in thought; see, e.g., Quine (1951, 1953). On the other hand, one might think that

Tarski showed how at least one commonsense notion of interpretation could be regimented and put to use in a fruitful way.[13]

Goodman offered a proposal for how to replace ordinary notions of meaning with a substitute that was in one sense extensional—without insisting that each word has an extension that is the word's meaning, and hence that 'Hesperus' and 'Phosphorus' are synonymous, along with 'unicorn' and 'centaur'. For Goodman, it was clear that 'unicorn' and 'centaur' have different meanings. The challenge was to defend this truism from skeptical attack without relying on intuitions regarding a vague and interest-relative (or objectionably mentalistic) notion of meaning. The idea was that even a tough-minded physicalist can agree that there are pictures of unicorns, and spatially distinct pictures of centaurs, and hence that the class of unicorn-pictures differs from the class of centaur-pictures. Similarly, Goodman claimed, one need not posit any mythical beasts to distinguish descriptions of unicorns from descriptions of centaurs. But if 'depiction of a unicorn' and 'depiction of a centaur' have different extensions, one can say that 'unicorn' and 'centaur' have different reference conditions—despite having the same (null) extension—because they appear as *constituents* of expressions that have different extensions.

This suggestion exploits a familiar idea in reverse. One might think that the phrases 'creature with a heart' and 'creature with a kidney' have different meanings because they have constituents that have different meanings. Even an extensionalist can agree that 'heart' and 'kidney' differ semantically. Of course, a certain kind of extensionalist might also think that the extension of a phrase is determined by the extensions of the constituent words, given the phrasal syntax. But one can relax this "bottom up" conception of how reference conditions are determined, in favor of a more holistic conception, as suggested by Frege's (1892a; 1892b) notion of abstraction and his claim that words signify things only in sentential contexts.

Suppose you learn that 'picture of a unicorn' and 'picture of a centaur' have different extensions. If you assume that meanings determine extensions, you can conclude that 'unicorn' and 'centaur' have different meanings, and use what you've learned in figuring out what the words mean. This doesn't imply that meaning is reference. But if two expressions are extensionally equivalent across

13 Davidson (1967) offered a partly reconstructive proposal that explicitly tied linguistic meaning to Tarski's notion of truth. This sparked objections and illuminating discussions; see, e.g., Davies (1981); Dummett (1976); Evans and McDowell (1976); Foster (1976); Harman (1974). But before this boom, there was even less clarity about a range of questions concerning whether meaningful expressions have meanings, how sentence meanings are related to word meanings, and how meaning is related to understanding, thought, truth, reference, use, etc.

all linguistic contexts (including 'description of a...') for which sameness of extension can be determined, then in an important sense, the expressions are semantically equivalent. In short, Goodman treated "global" sameness of reference as a proxy for sameness of meaning. Or abbreviating further, "same reference, same meaning."

However, Chomsky's appeal to Goodman figured as one part of a far more interesting proposal: a theory of understanding is a theory of syntax *supplemented* by an account of reference conditions for generable expressions. So in the end, Chomsky's (1957) proposed proxies for meanings were multi-leveled syntactic analyses together with specifications (not yet provided) of reference conditions. But he didn't assume that generable phrases have extensions that are determined by the extensions of their lexical constituents. And as discussed above, in the context of syncategoremata, he avoided assumptions about how meanings can combine. So especially in light of Chomsky's (1977; 1995b; 2000b) discussions of reference—including the many ways in which words like 'water', 'river', and 'London' can be used—we shouldn't read his early references to Goodman as suggesting that ordinary words have classical extensions that somehow determine truth conditions for sentences.[14]

Chomsky (1977) returned to these issues, discussing examples like (36) and (37).

(36) beavers are mammals

(37) beavers build dams

Intuitively, (36) implies that *all* beavers are mammals, while (37) implies only that *typical* beavers—in the wild, with suitable resources and able to express their beaverish nature—build dams. Moreover, (37) differs in meaning from the passive (38),

(38) dams are built by beavers

which falsely implies that typical dams are built by beavers; cp. (39) and the equivalent (40).

14 The conditions imposed by ordinary words on human acts of reference are context sensitive in many ways, as suggested by Wittgenstein (1953), Austin (1961, 1962) and Strawson (1950). See Travis (2008), McGilvray (1991, 1999), Stainton (2008), and Pietroski (2017c) for discussion. Though despite this complexity, grammar can impose formal constraints on coreference (i.e., anaphoric dependence); see, e.g., Chomsky (1986).

(39) beavers built this dam

(40) this dam was built by beavers

Or consider (41) and (42),

(41) poems are written by fools like me

(42) mountains are climbed by fools like me

which imply that all poems are written by fools, but not that all mountains are climbed by fools. While there are systematic form-meaning correspondences that reflect details of lexical meanings, the relevant notion(s) of composition remain unclear.

> Global properties of the sentence, which may be quite involved, seem to play a role. We cannot simply assign a meaning to the subject and a meaning to the predicate (or to a sentence form with a variable standing for the subject), and then combine the two. Rather, the meaning assigned to each phrase depends on the form of the phrase with which it is paired (Chomsky (1977, p, 31).

Chomsky (1957) was already stressing that correlations between grammatical forms and stable features of the environment (including language-independent thoughts) are imperfect. But the "correspondences between formal and semantic features cannot be ignored." They should be studied "in some more general theory of language that will include a theory of linguistic form and a theory of the use of language as subparts (p. 102)." Indeed, he concludes *Syntactic Structures* with a cautiously optimistic remark about semantics.

> Description of meaning can profitably refer to this underlying syntactic framework, although systematic semantic considerations are apparently not helpful in determining it in the first place. The notion of "structural meaning" as opposed to "lexical meaning", however, appears to be quite suspect, and it is questionable that the grammatical devices available in language are used consistently enough so that meaning can be assigned to them directly. Nevertheless, we do find many important correlations, quite naturally, between syntactic structure and meaning; or, to put it differently, we find that the grammatical devices are used quite systematically. These correlations could form part of the subject matter for a more general theory of language concerned with syntax and semantics and their points of connection (p.108).

5 Recap

Chomsky (1957) offered a program for developing theories of understanding—and in that sense, theories of meaning for human languages—that combine formal theories of syntax with proposals regarding what competent speakers know about how expressions can/cannot be used, especially in acts of reference. He envisioned studies of form and meaning that developed in tandem. The phenomenon of constructional homonymity played a key role in formulating and testing his claims about how humans generate linguistic expressions. And in the final chapter, "Syntax and Semantics," he summarized the previous discussion as follows.

> We have now found cases of sentences that are understood in more than one way and are ambiguously represented on the transformational level (though not on other levels) and cases of sentences that are understood in a similar manner and are similarly represented on the transformational level alone. This gives an independent justification and motivation for description of language in terms of transformational structure, and for the establishment of transformational representation as a linguistic level with the same fundamental character as other levels. Furthermore it adds force to the suggestion that the process of "understanding a sentence" can be explained in part in terms of the notion of linguistic level. In particular, in order to understand a sentence it is necessary to know the kernel sentences from which it originates (more precisely, the terminal strings underlying these kernel sentences) and the phrase structure of each of these elementary components, as well as the transformational history of development of the given sentence from these kernel sentences (p.92).

From a contemporary perspective dominated by model theory and truth-theoretic conceptions of meaning, it can be tempting to view Chomsky's contributions to semantics more narrowly—i.e., as an initial specification of the expressions whose semantic properties are to be specified by some other algorithm that pairs sentences (relative to analyses) with representations of truth conditions. But Davidson's (1967) bold conjecture was never the only game in town. Chomsky (1957) had already offered a sophisticated proposal, motivated by his sense that

> There is no aspect of linguistic study more subject to confusion and more in need of clear and careful formulation than that which deals with the points of connection between syntax and semantics (p.93).

Bibliography

Austin, John. 1961. *Philosophical Papers*. Oxford, UK: Oxford University Press.

Austin, John. 1962. *How to Do Things with Words*. Oxford, UK: Oxford University Press.

Barker, Chris, and Pauline I. Jacobson. 2007. *Direct Compositionality*. Oxford, UK: Oxford University Press.

Berwick, Robert C, Paul Pietroski, Beracah Yankama, and Noam Chomsky. 2011. Poverty of the stimulus revisited. *Cognitive Science* 35:1207–1242.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague, The Netherlands: Mouton de Gruyter.

Chomsky, Noam. 1964. *Current Issues in Linguistic Theory*. The Hague: Mouton.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1977. *Essays on Form and Interpretation*. New York, NY: North Holland.

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

Chomsky, Noam. 1995a. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995b. Language and Nature. *Mind* 104:1–61.

Chomsky, Noam. 2000a. Minimalist Inquiries. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. Roger Andrew Martin, David Michaels, and Juan Uriagereka. MIT Press.

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

Chomsky, Noam. 2017. The Galilean Challenge. *Inference* 3.

Church, Alonzo. 1941. *The Calculi of Lambda Conversion*. Princeton, NJ: Princeton University Press.

Davidson, Donald. 1967. Truth and Meaning. *synthese* 17:304–323.

Davies, Martin. 1981. Meaning, structure and understanding. *Synthese* 48:135–161.

Dummett, Michael. 1976. What Is a Theory of Meaning? In *Truth and Meaning: Essays in Semantics*, ed. Gareth Evans and John McDowell. Oxford, UK: Clarendon.

Evans, Gareth, and John McDowell, ed. 1976. *Truth and Meaning*. Oxord, UK: Oxford University Press.

Foster, John. 1976. Meaning and Truth Theory. In *Truth and Meaning*, ed. Gareth Evans and John McDowell. Oxord, UK: Oxford University Press.

Frege, Gottlob. 1892a. Function and Concept. In *Translations from the Philosophical Writings of Gottlob Frege*, ed. Peter Geach and Max Black. Oxford, UK: Blackwell.

Frege, Gottlob. 1892b. On Sense and Reference. In *Translations from the Philosophical Writings of Gottlob Frege*, ed. Peter Geach and Max Black. Oxford, UK: Blackwell.

Goodman, Nelson. 1949. On Likeness of Meaning. *Analysis* 10:1–7.

Goodman, Nelson. 1951. *The Structure of Appearance*. Cambridge, MA: Harvard University Press.

Goodman, Nelson. 1953. On some differences about meaning. *Analysis* 13:90–96.

Harman, Gilbert. 1974. Meaning and Semantics. In *Semantics and Philosophy*, ed. K. Munitz and P. Unger, 1–16. New York, NY: New York University Press.

Higginbotham, James. 1983. The logic of perceptual reports: An extensional alternative to situation semantics. *The Journal of Philosophy* 80:100–127.

Higginbotham, James. 1985. On Semantics. *Linguistic inquiry* 16:547–593.

Horn, Laurence. 1989. *A natural history of negation*. Chicago, IL: Chicago University Press.

Idsardi, William. forthcoming. Why is Phonology Different? No Recursion. In Martin and Gallego, eds.

Katz, Jerrold, and Paul Martin Postal. 1964. *An Integrated Theory of Lingquistic Descriptions*. Cambridge, MA: MIT Press.

Katz, Jerrold J., and Jerry A. Fodor. 1963. The structure of a semantic theory. *Language* 39:170–210.

Laka, Itziar. 1994. *On the syntax of negation*. New York, NY: Garland.

Lasnik, Howard. 1999. *Minimalist Analysis*. Oxford, UK: Blackwell.

Lasnik, Howard, and Joseph J Kupin. 1977. A restrictive theory of transformational grammar. *Theoretical Linguistics* 4:173–196.

Lepore, Ernest. 1983. What Model-Theoretic Semantics Cannot Do. *Synthese* 54:167–187.

Lewis, David. 1970. General semantics. *Synthese* 22:18–67.

Lewis, David. 1975. Languages and Language. In *Minnesota Studies in the Philosophy of Science*, ed. K. Gunderson, volume 7. Minneapolis, MN: University of Minnesota Press.

May, Robert. 1985. *Logical Form: Its structure and derivation*. Cambridge, MA: MIT Press.

McGilvray, James A. 1991. *Tense, Reference, and Worldmaking*. McGill-Queen's Press.

McGilvray, James A. 1999. *Chomsky: Language, Mind and Politics*. Cambridge, UK: Polity Press.

Montague, Richard. 1970. Universal grammar. *Theoria* 36:373–398.

Pesetsky, David. 1995. *Zero Syntax*. Cambridge, MA: MIT Press.

Pietroski, Paul. 2017abc. Semantic Internalism. In *The Cambridge Companion to Chomsky*, ed. J McGilvray, 196–216. Cambridge, UK: Cambridge University Press.

Pietroski, Paul. 2017b. Semantic Typology and Composition. In *The Science of Meaning*, ed. B. Rabern and D. Ball. Oxford, UK: Oxford University Press.

Pietroski, Paul. 2017c. *Conjoining Meanings: Semantics without Truth Values*. Oxford, UK: Oxford University Press.

Pietroski, Paul. forthcoming. Limiting Semantic Types. In Martin and Gallego, eds.

Quine, Willard Van Orman. 1951. Two Dogmas of Empiricism. *Philosophical Review* 60:20–43.

Quine, Willard Van Orman. 1953. *From a Logical Point of View*. Cambridge, MA: Harvard University Press.

Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral Dissertation, MIT, Cambridge, MA. Published as *Infinite Syntax!* Norwood, N.J.: Ablex (1986).

Russell, Bertrand. 1905. On denoting. *Mind* 14:479–93.

Russell, Bertrand. 1912. *The Problems of Philosophy*. London, UK: Butterworth.

Russell, Bertrand. 1957. Mr. Strawson on Referring. *Mind* 66:385–89.

Stainton, Robert. 2008. Meaning and Reference: Some Chomskyan Themes. In *Oxford Handbook of Philosophy of Language*, ed. Ernest Lepore and Barry C. Smith. Oxford, UK: Oxford University Press.

Strawson, Peter F. 1950. On referring. *Mind* 59:320–344.

Tarski, Alfred. 1933. The Concept of Truth in Formalized Languages. In *Logic, Semantics and Metamathematics*, ed. John. Corcoran. Indianapolis, IN: Hackett, 2 edition.

Tarski, Alfred. 1944. The Semantic Conception of Truth. *Philosophy and Phenomenological Research* 4:341–75.

Travis, Charles. 2008. *Occasion-sensitivity: Selected essays*. Oxford, UK: Oxford University Press.

Turing, Alan. 1936. On Computable Numbers with an Application to the Entscheidungs problem. In *Proceedings of the London Mathematical Society 2*, volume 42, 230–265.

Wittgenstein, Ludwig. 1953. *Philosophical Investigations*. New York, NY: MacMillan.