January 2008

Semantic universals and typology

Emmon W. Bach
University of Massachusetts - Amherst, ebach@linguist.umass.edu

Wynn Chao
School of Oriental and African Studies, wynnbach@gmail.com

Follow this and additional works at: https://scholarworks.umass.edu/linguist_faculty_pubs

Recommended Citation
Retrieved from https://scholarworks.umass.edu/linguist_faculty_pubs/1

This is brought to you for free and open access by the Linguistics at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Linguistics Department Faculty Publication Series by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
ON SEMANTIC UNIVERSALS AND TYPOLOGY

Emmon Bach
Wynn Chao

0.0. Language universals and typology

At the time of the Dobbs Ferry conference on Language Universals (1961), which culminated in the publication of *Universals of Language* (Greenberg, 1963/1966), American linguistics was in the midst of a transition from a strongly empiricist stance to the more rationalist approach characteristic of generative grammar and various subsequent developments. Perhaps the most influential statement of the empiricist view of the preceding decades is embodied in Bloomfield’s (1933) *Language*. The structuralist approach to linguistic research, with its emphasis on rigorous observation and description, is eloquently expressed in Bloomfield’s famous dictum, referred to in many of the papers in the *Universals* volume:

> The only useful generalizations about language are inductive generalizations. (Bloomfield, 1933: 20, in Greenberg, 1963/1966: 1, 67, 218, 281, 303)

For Bloomfield, ‘those areas of language study unamenable to such rigorous discipline were simply abandoned or relegated to the periphery: psycholinguistics, philosophy of language, and much of semantics.’ (Robins, 1988: 481)

This prevailing view received a strong challenge with the publication of Chomsky’s (1957) *Syntactic Structures*, with its rationalist view that fundamental aspects of language knowledge were to be explained in terms of a universal grammar in the speakers’ minds.

This difference was only the most recent manifestation of a longstanding conflict between the rationalist and empiricist approaches to linguistics, which first appears in the arguments between the ‘rationalist’ Stoic (300 BC to 529 AD) and the Alexandrian grammarians' views on language. With some exceptions, the didactic, descriptive *grammatica civilis* approach predominated until the eleventh century, when the revival of Greek scholarship and the influence of Islamic thought led to the development of the approach described by Campanella as *grammatica philosophica*, which was primarily concerned with ‘relating a descriptive framework to a theory of language’ (Robins 1988: 463-70)

The seventeenth century Port Royal grammarians, among the main exponents of the philosophical school, explicitly set out to explain universal aspects of language in terms of logical and semantic factors in the speakers’ minds (*op. cit.* p. 477). Their rationalist approach is criticized centuries later in Bloomfield (1933:6-7), who urges us ‘to return to the problem of general grammar’ (*op. cit.* p. 20) only after enough data from enough languages is collected to allow for ‘not speculative, but inductive’ explanations.
In retrospect, the 1961 conference in language universals and its resulting volume marked one of the pivotal events in this 20th century engagement between the dominant empiricist and the emerging rationalist-theoretical paradigm. Its legacy endures to this day. Among other influential contributions, Greenberg's paper on universals and the basic word order typology (Greenberg, 1963a) practically spawned (or reinvigorated) a whole subfield of research on language universals and variation. Typological studies initiated by his work have provided a rich mine for linguistic theories, forming the empirical basis for much subsequent research on linguistic universals and universal grammar.

In the Greenberg volume, Charles Hockett discussed the question of the validity of crosslinguistic identification of lexical categories:

It was at one time assumed that all languages distinguish between nouns and verbs -- by some suitable and sufficiently formal definition of those terms. One form of this assumption is that all languages have two distinct types of stems (in addition, possibly, to various other types), which by virtue of their behaviour in inflection (if any) and in syntax can appropriately be labeled nouns and verbs. (Hockett:1963/1966 p. 4).

He went on to say that this claim had been invalidated by Nootka [= Nuu-chah-nulth] at the level of stems, but upheld in the higher categories of the syntax.

Hockett acknowledged that ‘(1.1) The assertion of a language universal must be founded on extrapolation as well as empirical evidence’ (p. 2), and ‘(1.10) The problem of language universals is not independent of our choice of assumptions and methodology in analyzing single languages’ (p. 7). Hockett’s assessment of the Port Royal grammar is significantly more positive than Bloomfield’s (Hockett, 1963/1966, pp. 4-5):

The Port Royal Grammar constituted both a putative description of language universals and the basis of a taxonomy. The underlying assumption was that every language must provide, by one means or another, for all points in the grammatico-logical scheme described in the Grammar. Latin, of course, stood at the origin in this particular coordinate system. Any other language could be characterized typologically by listing the ways in which its machinery for satisfying the universal scheme deviated from that of Latin. This classical view in general grammar and its taxonomy has been set aside not because it is false in some logical sense but because it has proved clumsy for many languages: it tends to conceal differences that we have come to believe are important, and to reveal some that we now think are trivial.

Hockett’s disagreement was thus on empirical, not theoretical grounds. Interestingly, the Port Royal Grammar used the exact form of argument that we see in much contemporary work. For example, it appealed to universal categories like Participle, Verb, and Adjective and their typical configurational properties to explain the distribution of agreement facts in the syntax of French. In fact, as pointed out in Bach, (1965a, 1971) Greenberg's and subsequent proposals about syntactic typology and universals could not be stated without the crucial assumption that syntactic categories across languages can be identified or compared and the proposals formed a basis for testing linguistic theories.

In this chapter we take Greenberg's paper as our inspiration. Starting from the assumption that syntactic categories can be universally identified or correlated (whether or not they are
in every languages) we wish to investigate the relation between them and their semantic interpretations, focussing on the nominal domain. Note that there are two possibilities here: first, there might be a universal stock of categories from which individual languages might draw; second, there may be hypotheses that all languages must instantiate particular categories. The situation in phonological systems is illuminating: the stock of possible sounds is given by a general theory of phonetics-phonology, but not all of the categories need be utilized in every language.

0.1. Semantic universals

In another paper in the volume entitled "On the semantic structure of language" Uriel Weinreich began by citing two assumptions that would be agreed on by most linguists. One of them sounds like an endorsement of some version of the so-called Sapir-Whorf hypothesis:

The semantic mapping of the universe by a language is, in general, arbitrary, and the semantic "map" of each language is different from those of all other languages. (in Greenberg, 1963/1966: 142 - 216).

Interestingly, Weinreich goes on in the same article to give a quite detailed account of what might fairly be called universals of semantics. So it would seem that rather than denying that there is any uniform 'map of the universe' across languages he is actually looking for the proper apportionment of parts of the map to the universal and the particular.

One enduring distinction brought out in Weinreich's paper as well as in the papers by Ullmann and Jakobson in the Greenberg volume is the distinction between grammatical/structural and lexical meanings. We take this up below and suggest that the distinction may play a crucial role in helping us to reconcile a supposed contradiction between model-theoretic and conceptual approaches to semantics.

One real advance since Dobbs Ferry comes from the logical and philosophical ideas that have become part of the semanticist's toolbox. At the time of Dobbs Ferry, the main logical tools that linguists were aware of were those of first order logic and this generally went together with a syntactic view of interpretation. In the following decade and a half a much richer set of possibilities was opened up. The work of Richard Montague was central and we will make use of his work and the tradition that followed from it here. Somewhat surprisingly, the model-theoretic approach7 of Montague and his followers made possible a much closer fit between language and interpretation than in much previous and subsequent work in other traditions.8

Example: In the logical tradition going back to Bertrand Russell, it was customary to explicate the meaning of a definite description in a sentence like the king of France is bald with a logical form that split the meaning of the subject into three pieces that do not correspond to any constituent of the English sentence:

$$\exists x[\text{KingFrance}(x) \& \forall y[\text{KingFrance}(y) \leftrightarrow y = x]]$$

where we have an existential quantifier with a variable, an association of the variable with the predicate King of France, and a uniqueness clause (if anything is a King of France it is identical with the entity picked out by the variable). In Montague’s most widely read and cited work on English (1973: PTQ), all term-phrases are assigned to a
single syntactic category with a uniform semantic interpretation as **generalized quantifiers** (more on this below).

In the main, we follow the model-theoretic program in this paper.

**Syntax-semantics mapping**

Is there a uniform way of mapping linguistic expressions of various sorts to interpretations? We assume that the best answer to this would be **Yes**. If we are asking this question across languages again the best answer would be **Yes**. Notice that the second question doesn't make much sense unless we can identify or relate syntactic categories across languages.

This answer relies on a methodological strategy, we may call “Shoot for the Universal”, implicit in a lot of linguistic theorizing (explicitly invoked in Bach 1968; Hankamer 1971). The reasoning behind these strategies is that the best way to find out about a domain is to make strong claims and retreat from them only in the face of contradictory evidence. Stronger claims make for more detailed predictions.

In order to talk about these mappings we need to say something about the structure of the models we are considering.

There are three kinds of questions we can ask about semantic universals and typology from the point of view of model theoretic semantics

(a) Are the basic elements of the model structure universal?
(b) Are the relations between the syntactic categories and semantic interpretations universal?
(c) Are there typological patternings related to either of (a) or (b)?

The issues discussed in the foregoing parts of our chapter, reflect an ongoing tension in linguistics between description and theory (Bach 2004). In our view, there should be a fruitful complementarity between theoretically informed description and empirically informed theorizing. Typological study is the natural meeting ground between these two activities.

**0.2 Syntax and Semantics**

Here's how we proceed. A language is a pairing of expressions and meanings, among other things perhaps. We describe the language by means of an explicit or generative grammar with a lexicon. The grammar tells us about the form of these complex expressions as well as about the meanings. The form and the meaning of the complex expressions are dependent on the form and meaning of the expressions that are combined. The step by step or locally confined process of assigning form and meaning to a resultant expression as a function of the form and meaning of the input expressions conforms to the requirement of **compositionality** (see below).

What do we mean by a meaning? In the model theoretic view, a meaning is something that is not language, something that linguistic expressions refer to, or **denotation**. The grammar assigns denotations to linguistic expressions. The whole system of possible denotations makes up a **model structure**.

We hypothesize that the general model structure is the same for all natural languages. It is very simple.

First of all, since some linguistic expressions like names can refer to individuals, we assume that there is a set of entities to which such names can apply and which is available for making
general statements by quantification or other means. Semantic theory places no restrictions on what can be an entity, as long as we don't get into any logical problems. In Montague's most often quoted paper on natural language (Montague, 1973), the entities of the grammar are the individuals which correspond to the names John, Mary, Bill, and ninety (as well as potential values for an infinite set of variables). We think it is safe to assume that every language provides for proper names. (How they might be constructed is another matter.) Second, we want to say that sentences can be true or false, so we include two truth values.

Third, as a value for predicates like laughs we want to have something like a set of entities: the entities that laugh, in this case. A way to model this is to think that a predicate like the denotation of laughs is a function from entities to truth values:

(1) Guinevere laughs.

This functional way of looking at semantic expressions that are more complex than simple names is generalized to model all sorts of denotations as functions. Moreover, they can all be modelled as unary functions: taking one thing to give a value, since the value itself can be a function. So for example, to model simple denotations of transitive verbs, we take them to denote functions from arguments to intransitive verb denotations -- that is, to predicates. Here's an example:

(2) Guinevere ignores Lancelot.

The phrase ignores Lancelot denotes a function from entities to truth values, while ignores itself denotes a function from entities to predicates or intransitive verb phrase denotations. This choice generalizes to give two-part constructions (binarism) as the general pattern.

A primary benefit of adopting a functional view (in the mathematical sense of function) is that we can have all the properties of functions without special stipulation. Here for example we can model two place relations or functions like those associated with transitive verbs by a step-wise analysis into unary functions (to functions etc). This process is known as currying. In this case the step is well-motivated linguistically. A second free benefit is the possibility of combining functions by composition.

One further step is required before we can go on. The names in the examples given so far can be replaced (in English) by a wider set of expressions like every knight, some queen, the giant in the castle. Montague introduced the idea that such expressions could be interpreted as generalized quantifiers, that is, sets of sets or sets of properties. He assimilated proper names to the same type. So in sentences like the ones used so far, the expression Guinevere is taken to denote the set of all Guinevere's properties, including the properties of laughing, ignoring Lancelot, and so on. This move had far-reaching consequences, some of which we mention below.

So far we have a model structure that is appropriate for interpreting sentences that deal with the real world. But natural languages allow us to talk not only about the way things are, but about the way things might be:

(3) Guinevere might love Arthur.
We understand this sentence to mean that it is not excluded that Guinevere loves Lancelot either as things are or if circumstances were different. The way in which this will be modeled is to introduce the notion of ways that things are or may be. In Montague and related approaches this leads to the assumption that there is a set of possible worlds to help interpret sentences like (3) as well as other expressions, sometimes instead of or in addition to a set of situations, world-like ways that things might be, but of a "smaller" kind.

Including possible worlds or situations brings with it the possibility of modeling other special kinds of meanings. For example, a function from possible worlds to individuals gives us what Montague called an individual concept. A function from possible worlds to sets or predicates gives us one way of thinking about properties, ways of finding the sets of entities that are instances of the predicate in any world. We say more about these intensional objects below.

In order to talk about this we need to say something about the structure of the model beyond what we have done so far.

We associate each of the ingredients of the interpretation with a type.

\[
\begin{align*}
\text{t} & \text{ is the type for truth values} \\
\text{s} & \text{ is the type of situations or worlds} \\
\text{e} & \text{ is the type for entities}
\end{align*}
\]

So we can model the structure of denotations, using these ingredients and notating functional type like this: if \(a\) and \(b\) are types then there is a type of functions from \(a\)-type things to \(b\)-type things which we represent like this: \(<a, b>\). This way of looking at the structure of denotations forces complex meanings to be built up pairwise. So the type for transitive verbs (with extensional interpretation) is \(<e, <e, t>>\), as illustrated in Example (2).11

**Compositionality**

The principle of compositionality requires that the meaning of a complex expression must be a function of the meaning of the parts and the way in which they are combined.

This requirement is generally appealed to in some form or other, but needs explication (Partee 2004). In a recent paper David Dowty (2007) argues that one can't talk about the problem of semantic compositionality without thinking of syntactic options: tighter vs. greater degrees of freedom in syntax, strict vs. extended categorical grammar, in Montague’s general theory these would be the allowed syntactic operations.

So much for a barest sketch of the general denotational space of our semantics. We now turn to some examples and implications for semantic universals and typology.

**0.3. Elaborating the domain of entities**

The general model structure outlined above is not enough to model the distinctions necessary for a good account of natural language semantics. We will look at a number of elaborations
of the model structures that have been proposed, but first let us show the limitations of trying to build meanings just out of the ingredients of the model structure as given so far.

Early and late, people have used this basic model structure to construct various kinds of higher order denotations. A prominent move has been the introduction of various kinds of intensional entities, a move which began (in modern times) with Frege and continued with Montague's own ventures into natural language semantics. Thus to solve various kinds of puzzles and problems, Montague used individual concepts, properties, and propositions, all of these being functions from worlds (or world-time pairs in PTQ) to various entities and sets. The distinction between intensions and extensions is one way of reconstructing Frege’s distinction between sense and reference (Sinn and Bedeutung: Frege, 1892).

**Example:** it may be that the set of two-legged rational animals and the set of humans are the same set in this world. But there are surely worlds where they are not, so we can say that the property of being a two-legged rational animal and the property of being human are not the same. Similarly for individual concepts like those named by the Morning Star and the Evening Star both of which refer to the planet Venus.

Often these constructs have not been as finegrained as needed and various of them have been introduced into the model structures as independent and primitive elements: propositions and properties, for example (Thomason, 1980; Chierchia, 1984).

The principal elaboration that we will follow here goes by way of dividing up the set of entities A into Sorts. A sort is a distinguished subset of the domain which allows us to make finer discriminations than is offered by the basic model structure. We mention a few sorts that will be used below in our discussion of semantic typlogies.

Greg Carlson (1977) introduced Kinds, (ordinary) Individuals, and Stages into the model, all as Sorts of elements in the domain of individuals but linked by relations such as that between a Kind and instances of the Kind, and between Individuals and Stages, Stages being something like temporal slices of the manifestation of an individual in a history or world or situation.

Chierchia (1984) added Properties as independent primitive elements. Chierchia made use of operators that make predicates from properties and properties from predicates, in effect. He also marshalled a considerable amount of evidence for the independence and necessity of treating properties as basic elements.

Plural and Mass entities were introduced in Link (1983). Plural objects can be freely formed from the set of singular entities and used to model the denotations of words like "dogs": the denotation of singular "dog" is the set of individual or atomic dogs, "dogs" denotes, among other things, the whole set of dogs as a plural entity as well as all pairs, triples, and so on, such as the three dogs in my house -- Fido, Caesar, and Pompey, for example --, but minus the atoms. This algebra for modeling the meanings of count nouns was matched by a non-atomic algebra for mass terms and the "stuff" of all entities.

We assume Linkian kinds of structures for the denotations of nouns (Link, 1983; see also Scha, 1981; Landman, 2000), as mentioned. For simplicity we use singleton sets in place of atoms but we will still refer to ‘singulary’ elements in the structures as "atoms" (following Rullmann and You, 2006).
**Example:** Suppose there are three dogs a, b, c in a situation or world. The whole structure looks like this:

\[
\{a, b, c\}
\]
\[
\{a, b\} \quad \{a, c\} \quad \{b, c\}
\]
\[
\{a\} \quad \{b\} \quad \{c\}
\]

Let's refer to the interpretations for these various elements as **singularities** (or **atoms**), below the line here, **pluralities** above the line, and **transpluralities** for the whole domain. We don't take any position on whether the atoms are to be treated as singleton sets as suggested by the picture, or as atoms **simpliciter**. The null set is excluded so that predications about dogs, for example, would not be trivially satisfied in dogless worlds.

For English, we can adopt a straightforward association of the various domains of plurality with the different syntactic expressions. As in Link (1983), singular common nouns are taken to denote sets of singular individuals, and plurals to denote sets of pluralities. The union of the two sets of denotations is an appropriate place for expressions of "general number" (Corbett, 2000, and see the discussion in the next section).

Taking these suggestions together, we suppose that the domain A of entities has at least these sorts: (ordinary) Individuals, Kinds, various sorts of individuals that can be modelled in the domains for Plurals, Masses, etc. Kinds can be related by a taxonomic relation ("subspecies of", see Krifka, 1995), while Kinds can be instantiated in Individuals by a realization relation (Carlson, 1977; Krifka, 1995).

### 0.4. Some semantic typologies

Logically speaking, typologies require variety. If some property of languages is truly universal, then all languages will be of the same type as far as that property is concerned. So typological investigations begin by noting or proposing some characteristic that is not universal. Interesting results come when it is possible to see clusterings of properties.

One such area of investigation came directly out of the modeling of the interpretations for term phrases (DP's) as generalized quantifiers, mentioned above. In the course of their detailed study of generalized quantifiers in natural languages, Barwise and Cooper (1981) enunciated the following hypothesis:

**NP-Quantifier universal:** *Every natural language has syntactic constituents (called noun-phrases) whose semantic function is to express generalized quantifiers over the domain of discourse.* [Note that NP here corresponds to what is now usually called DP following Abney 1987]
Research on a number of languages as well as a closer investigation of the possibility of other ways of expressing quantification led to a classification of quantification strategies into so-called **D-Quantification** and **A-Quantification** -- see below (Bach et al. 1995 Introduction and Partee's contribution to that volume: Partee 1995). This classification has immediate typological consequences, of which we mention one cluster that was made prominent in the work of Eloise Jelinek (1984, 1995):

**Quantification: A and D quantification**

D-Quantification is the strategy whereby generalized quantifiers are available as the denotations of term-phrases (DP's: 4a) and is opposed to A-quantification by means of Adverbs (and other means which coincidentally are associated with categories beginning with "A": Auxiliaries and Affixes), as in examples made prominent by David Lewis (1975: 4b), and brought into linguistics especially in the theories of Irene Heim (1983) and Hans Kamp (1981):

(4a) *Every commuter will read a newspaper.*

(4b) *Usually, a commuter will read a newspaper.*

Lewis showed that examples like 4b could not be resolved into the usual sort of quantification on individual variables, but had to be thought of as quantification of instances or cases involving several variables. Adverbs like *usually* in (4b) were treated as **unselective quantifiers**.

Suppose now that there are languages without A-quantification or without D-quantification, then we might have a basis for a semantic typology: languages with only one or the other or both. These differences might reflect independent parameters or be consequences of some other property (see discussion of pronominal arguments below).

It has been claimed that there are languages that have no quantification at all, a claim that has led to a widely publicized controversy: see Everett 2005 on Pirahã, critique in Nevins et al. 2007, and reply by Everett 2007.

Investigations and discussions like those just mentioned involve in an essential way both theory and description. The putative universal proposed by Barwise and Cooper (1981, see above) requires specific claims about syntactic structures, semantic interpretations, and the relations between them.

**Pronominal argument languages and generalized quantifiers**

Jelinek (1984) proposed the **pronominal argument hypothesis** originally to account for differences of configurationality. According to this hypothesis some languages require that verbs have affixed or cliticized pronouns as their arguments, and not full DP's. Semantically, this would mean that the verbs are in effect to be interpreted as something like open sentences, with unbound variables in the argument slots. Jelinek's hypothesis predicted several characteristics of such languages: free nominal expressions in sentences as something like adjuncts, relatively free word order, and (Jelinek 1995) the lack of quantificational DP's. In this second paper, Jelinek drew a further typological consequence. Languages without a noun-verb dichotomy would lack DP's as such and hence would have no generalized quantifier constituents.
Here is a schematic comparison of a sentence like John walks as it might be rendered in a language like English and one like Straits Salish (see Bach 1994: 274):

(5a) “English”: “Straits Salish”:
J’ (walk’) walk’(x) & john’(x)
Here J’ stands for the generalized quantifier: the set of all John’s properties, and the formula is interpreted as saying that this set includes the property of walking. In the pseudo-Straits-Salish example the best paraphrase is something like: he walks and he is John. No such paraphrase is possible with a formula that would correspond to Every man walks in a language with no D-quantification. (5b) shows the difference in expressing the generalization in the two strategies. The first expression shows that every man can be interpreted in a way exactly parallel to John but in the second instance, with A-quantification there is no unitary generalized quantifier interpretation:

(5b) [EVERY(man')](walk') ALWAYS[walk'(x), man'(x)]
(ALWAYS: unselective binder)

Nouns and verbs

The pronominal argument hypothesis is related to the question about the universality of nouns and verbs as lexical or syntactic categories. Swadesh (1939: 78-9) argues that in Nootka [=Nuu-chah-nulth] all stems can potentially be inflected as predicates; the distinction between N and V interpretations may not be present in the lexical domain, but resurfaces at the level of syntax, where the expression associated with verbal morphology and interpreted as the main predicate is always clause-initial:

(6) mamo·k-ma qoʔas-ʔi
work-3SG IND man-DET
‘The man is working.’

(7) qoʔas-ma mamo·k-ʔi
man-3SG IND work-DET
‘The working one is a man.’

This pattern is characteristic of Northwest Coast languages in the Wakashan and Salishan families, as these examples show:

Straits Salish (Jelinek, 1995: 490)

(8) swi’qo’əł + 0 ce t’iləm + la’
young man +3ABS DET sing +PAST
He is a young man, the one who sang.
The one who sang is a young man.

(9) t’iləm + la’+ 0 ce swi’qo’əł
sing + PAST +3ABS DET young man
He sang, the one who is a young man.
The young man sang.
Kwak'ala (retranscribed, Boas 1947: 280, Boas cites other languages here as well):

(10)  **N’ikida bəgwənəm.** `That one said, it was the man'
    n’ik- -ida bəgwənəm
    say- -infl  person

(11)  **Bəgwənəmida n’ika.** `It was the man he [i.e. who] said'
    bəgwənəm- -ida n’ik- -a
    person- infl  say- -completive

Haisla (EB fieldnotes)

(12)  **guxw** 'house' / '(to be a) house'

(13)  **Guxw gada.** `this is a house'
    house this

(14)  **Duqwelan qix guxwgaxhga.** `I see this house'
    duqwela-n qix guxw-gaxhga.
    see-I this house-here

The noun-verb question continues to cause discussion, see for example, Demirdache and Matthewson 1995, Evans and Osada, 2005

**Kinds and Plurality**

In the last few years there has been quite a lot of research into variation across languages with respect to nominal structures and their interpretations. Considerable attention has been paid to Mandarin Chinese (Krifka 1995, Chierchia 1998a, 1998b, Cheng and Sybesma, 1999; Rullmann and You 2006, among others).

There are three striking differences between English and Mandarin in the realm of nominal expressions.

i. First, all nouns in Mandarin can occur bare.

ii. Second, there is no expression of a singular – plural distinction in the morphology of the noun.12

iii. Third, nouns cannot be construed with numerals in Mandarin without the help of a classifier or measure expression.

Because these three properties are matched exactly in English by mass nouns like *blood, mud, or furniture*, Mandarin nouns are often claimed to be mass nouns in general. Note that English also has uncountable plurals such as *cattle, police, poultry, three head of cattle, three cattle*, (see Huddleston and Pullum, 2002: 345).

A variant of this idea is Chierchia’s view that mass nouns are basically plurals, which then leads to the claim that Mandarin nouns are also all basically plural (Chierchia, 1999b).
Mandarin nouns can be interpreted in various ways, depending on the syntax: definite, indefinite, or as names for Kinds (see Cheng and Sybesma, 1999, Chao and Bach 2004, and Rullmann and You 2006, for details):

(15)  
gǒu  
`dog, dogs, a dog, the dog, Dogs [generic]' 

(16)  
gǒu yào guò mālù 
dog want cross road 
`The [*a] dog wants to cross the road.' (Cheng and Sybesma, 1999)

(17)  
gǒu ài chī ròu 
dog like eat meat 
`Dogs love to eat meat.' (ibid.)

(18)  
wǒ kànjiàn xióng le  
I see bear aspect 
`I saw (some) bears.' (Krifka, 1995)

Chierchia (1999b) proposed a set of choices for NP interpretations, which led him to a typology of languages in their nominal systems. These choices were based on two features: arg and pred, each with plus [+] and minus [-] values. If the value for arg is plus for a language, then NP's can occur as arguments (of verbs), if minus then not. If the value for pred is plus then NP's can act as predicates (restrictors on quantificational DP's), if minus not. Having minus for both properties is excluded in principle. This setup leads to predictions for three kinds of languages, as exemplified thus:

NP[+arg,-pred]  nouns refer to kinds, every NP is of type e  
Mandarin, Japanese

NP[-arg, +pred] every noun is of type <e,t>: no bare nominals at all  
French  
Italian null D's only if licensed by lexical head

NP[+arg, +pred] NP's can be freely argumental or predicative  
English and Germanic languages

There have been critical reviews of this typology, see especially Cheng and Sybesma1999; Krifka 2004; Borer 2005a; Schmitt and Munn 1999; Rullmann and You 2006. Some of these discussions have brought in new data from languages that go against Chierchia’s typological predictions. For example, Brazilian Portuguese (Schmitt and Munn, 1999) allows bare singulars, a number of languages with optional plurals go against Chierchia’s claims that classifiers and plural marking cannot go together. Some of the critiques have claimed that the relevant parameters should be sought in grammar rather than semantics (Schmitt and Munn, 1999).

Rullmann and You (2006) can be cited as a promising typological study on a different semantic base. In their theory, “…the crucial parameter [in this domain] does not involve kind reference, but number: in some languages the extension of morphologically unmarked count nouns includes only atoms, whereas in other languages it includes both atoms and pluralities,” that is, in our terminology “transpluralities.”

Rullman and You's theory leads to the prediction that in languages with `general number' (with the whole set of transpluralities as domain) a situation involving one or more entities
might in principle be described using either a plain or a plural form of a noun. In other words, plural marking will be optional or 'facultative' (Corbett 2000).

There are many open questions at present in this area of research. For our purposes here, we simply want to stress that the explicit model-theoretic semantics has made it possible to propose quite precise hypotheses about the syntax-semantic mapping and typological consequences, both for syntax and semantics.

0.5. Concepts and properties

We mentioned above the distinction between grammatical or structural aspects of language and lexical aspects, reflected for example in the distinction between functional categories and open-class lexical categories. We also noted that it may be reasonable to relate this distinction to the distinction between model-theoretic and conceptual approaches to semantics.

We noted also that nominals of various kinds lead a double role, as shown in the last section and as discussed by various writers. On the one hand they are closely related to Kinds and on the other to predicates. Krifka (2004) argues that nouns are to be interpreted in the first place as Properties, and that their double rôle can be derived from this common base.

This idea fits well with conclusions reached by Bach in a series of papers on word-internal semantics, especially concerned with polysynthetic languages. After trying to come up with a reasonable choice for the basic denotations of elements involved in derivations of words -- both stems or roots and affixes -- he concludes that the only thing left might be Properties or perhaps nothing model-theoretic at all, perhaps Concepts in the psychological sense (see Bach 2005 and papers referred to there).

A prime instance is the textbook example of Semitic consonantal roots like k-t-b, which underlies a large number of derived lexical items and inflected forms. All the words are vaguely connected with the notion of writing, but it is difficult to see how they can be compositionally built up in a model-theoretic fashion. The same point can be made with English words derived from what is ultimately the Latin root scrib-: describe, inscription, conscription, scripture, and so on.

It is noteworthy that much of the rich literature on the analysis of word meanings by such writers as Jackendoff, Pustejovsky, and Talmy, is frankly based on conceptual foundations (see references under these names).

So it may be that Model Theoretic and Conceptual Semantics are complementary rather than competitors for the True Theory, the former suited to grammatical or syntactic structures, the latter to word-building and sublexical meanings. But there is no time or space to pursue this possibility here. In any case, it is worth pointing out again that Model-Theoretic semantics itself is compatible with a view that the denotations are to be sought in psychological models (Zwarts and Verkuyl 1994).

0.6. Some other syntactic-semantic investigations
We have touched here on only one area of the syntax-semantics complex. Similar investigations have been carried out in a number of other areas, and some have been related to each other. We simply mention them here, with little by way of elucidation:

Clause structures: Cinque's 'cartographic approach' maps out a large number of clausal domains, each with its functional projection and associated Specifiers and/or Modifiers (1999) and makes universal claims about the layering and ordering.

Nominal structures: We mentioned above only the inner and outer parts of the DP systems in several languages. The topic of adjectives and other nominal modifiers and other intermediate parts of the DP has been the topic of cross-linguistic investigation, in part following the ideas of Cinque just mentioned and as part of the investigation of wide-spread parallels in clausal and nominal structures (Scott 2002, 2003, Chao, Mui, and Scott [2001], Chao and Bach [2004]).

Verbal aspect: The syntax and semantics of aspect and aspectual verb classification (Aktionsarten) has been the subject of vigorous study since Verkuyl's and Dowty's pioneering work in the seventies (Verkuyl 1972, Dowty 1972, 1979). Distinctions such as those between events (in the narrow sense) and processes, states, achievements, accomplishments, and the like, deriving ultimately from Aristotle, have been modeled in a wide variety of ways, and parallels to the nominal distinctions between count and mass terms have been the locus of much discussion from both a formal and semantic point of view (Verkuyl et al. 2005).

If we add eventualities of various types as sorts we can accommodate special denotations for other categories. For example, it has been suggested that verbal constructions (or some verbal constructions) have an event argument (Kratzer 1995). The importance of events in the logic of natural language was insisted on by Donald Davidson (1967) long ago.

0.7. Outlook and conclusions

We began this essay with the delineation of two views of language universals as reflecting a tension between empiricist and rationalist views of the nature of scientific investigation. There will probably always be a tension between these two impulses, and not only in linguistics. We also suggested that typological studies are the natural meeting ground--or should we say "battle-ground"?--for these two impulses or stances.

We have probably shown our bias in this dichotomy by our choice of examples. Most of the studies cited work squarely within frameworks that treat universals as terms in hypothetico-deductive theories. But far from having no empirical import, most of the claims we have cited have led to fruitful new work on a wide variety of languages. In the last analysis, just as there can be no such thing as pure theory-free description, there cannot be interesting theories without confrontation with crucial empirical consequences.

We hope to have shown that a semantic perspective can lead to far reaching and interesting typological results and questions. The fact that our paper has been richer in questions than in firm results is a reflection of the vitality of the field.

1 An indication of the transitional position of the volume: there are 16 references to Bloomfield in the index to Greenberg, 1963/1966; 5 to Chomsky.
The term "general grammar" is used here in the sense of Arnauld and Launcelot's *Grammaire générale et raisonnée* (Anonymous, 1676 /[1966]).

The picture we give here needs considerable shading to be accurate: (1) structuralism cannot be equated with American structuralism, and (2) even in America, the influence of Bloomfield and his particular brand of empiricist behaviorism must not be exaggerated.

The debate about this claim for universality of the noun-verb distinction continues, compare Section 0.4 below and Evans and Osada (2005).

Compare Bloomfield 1933: p.6.

On lexical categories, see also Baker (2003), who makes a particularly strong argument for universal syntactic categories.

See below for an exposition of the basic assumptions of model-theoretic semantics. It is to be distinguished from conceptual semantics which seeks meaning in mental or psychological entities and to proof-theoretic semantics which models meanings in symbolic languages. A classic statement of the program of model-theoretic semantics is David Lewis’s 1972 paper, “General Semantics.”

Montague's advice: "Pay attention to natural languages, they may be trying to tell you something." Not his words, but often implicit in a lot of work in this tradition. Compare Montague's famous remark at the beginning of his paper "English as a formal language" (1970a): "I reject the contention that an important theoretical difference exists between formal and natural languages." Like PTQ this paper is about English and not some other language such as Logical Form that English can be mapped into.

The “other things” include aspects of meaning in the broad sense that are better handled in integrated but separate components dealing with presuppositions, pragmatics, implicatures and the like.

Other choices are possible, for example, a three-valued logic, with a third value being “undefined.” This elaboration is no doubt necessary if we adopt a model with Sorts as outlined below. For purposes of exposition, we stick to the classical system with two truth values.

Note that “function,” “functional” occur in at least two uses in linguistics: as a strictly mathematical term as in this paper, and in various syntactic frameworks as referring to “functional” categories as opposed to lexical or open-class categories. [The difference and relation between them is to be taken up in Bach and Chao, in preparation a].

We reject the claim that the Chinese suffix –men, which attaches to pronouns and some nouns, is an ordinary plural marker (see Li 1998). We hold that, like the similar Japanese suffix –tati, –men is a group-forming suffix, most closely corresponding to the English expression ‘_and them_’ as in _Joel and them_ meaning Joel and some group related to him by some contextually salient or conventional relation.
Key to further readings:
Here are some suggestions for further reading and background material:
A classic defense of model-theoretic semantic is Lewis 1972. Partee 1996 gives a
comprehensive survey. For a general introduction to conceptual semantics see Jackendoff
1996. Semantic typology is the focus of two papers by Chierchia (1998a, 1998b) with special
reference to the semantics of nominal expressions. For crosslinguistic studies on
quantification, see Bach et al. 1995. Barwise and Cooper 1981 presented the important NP-
universal hypothesis and provided a detailed study of the logic of generalized quantifiers.
Cinque 1999 is the basic reference to the cartographic view of clausal structures.
Dowty 1979; Levin and Rappaport Hovav 1996 are two basic works on lexical semantics.
Two recent works on compositionality are Partee 2004 and Dowty 2007. Link 1983 and
Landman 2000 are basic for mass terms, events, and plurality.

For additional special topics, check the references in the body of the paper and footnotes.

References:

MIT.

[Arnauld, Antoine, and Claude Lancelot]. Anonymous 1676/[1966]. Grammaire générale
et raisonnée ou La Grammaire de Port-Royal. Edition critique présenté par Herbert E.
Brekle. Facsimile of the third revised and expanded edition. Stuttgart-Bad Canstatt: Friedrich
Frommann Verlag (Guenther Holzboog). First edition: 1660.

University Monograph Series on Languages and Linguistics 18, 3-18.

Bach, Emmon. 1965b. Structural Linguistics and the Philosophy of Science. Diogenes
51, 111-128.


and P. Weingartner, eds. Logic, Methodology, and Philosophy of Science VII (Amsterdam et

John MacNamara and Gonzalo E. Reyes, eds., The Logical Foundations of Linguistic Theory


Evans, Nicholas and Toshiki Osada 2005 The myth of a language without word classes Linguistic Typology 9:351-390


Index
Abney, 16
Aktionsarten, 27
A-quantification, 17
arg, 22
Arnauld and Launcelot's Grammaire générale et raisonnée, 29
atoms, 14
Bach, 1, 4, 7, 16, 18, 25, 26, 29
Baker, 29
Baker, Mark. 2003, 33
Barwise, 16
Bloomfield, 1, 2, 29
Boas, 20
Borer, 23
Brazilian Portuguese, 23
Campanella, 2
Carlson, 14, 16
Carlson, Greg N. 1977, 34
Chao, 1, 26, 29
Cheng, 21, 22, 23
Chierchia, 13, 14, 21, 22
Chierchia, Gennaro. 1998a, 35
Chinese, 29
Chomsky, 2, 29
Cinque, 26
Clause structures, 26
compositionality, 8
Compositionality, 12
Concepts, 24, 25
Cooper, 16
Corbett, 24
currying, 10
Davidson, 27
definite description, 6
Demirdache, 18, 21
denotation, 8
denotational space, 12
Dobbs Ferry, 6
Dowty, 12, 27, 36
DP, 16
D-Quantification, 16
empiricist approaches to linguistics, 2
English, 18, 23
entity, 8
Evans, 21, 29
events, 27
eventualities, 27
Everett, 17
Everett, Daniel L. 2005, 37
extensions, 13
Frege, 13
French, 23
function, 29
functiona, 29
functional, 29
functional categories, 24
functions, 9
general grammar, 29
general number, 24
generalized quantifiers, 6, 10, 16
Germanic languages, 23
grammatica philosophica, 2
grammatical/structural and lexical meanings, 5
Greenberg, 1, 2, 4, 5, 29, 37
Haisla, 20
Hankamer, 7
Heim, 17
Hockett, 3
Huddleston, 21
individual concept, 11
individuals, 8
Individuals, 14
inductive generalizations, 1
inductive" explanations, 2
intensional entities, 13
intensional objects, 11
intensions, 13
Italian, 23
Jackendoff, 25
Jackendoff, Ray. 1996., 37
Jakobson, 5
Japanese, 23, 29
Jelinek, 17, 18, 19
Kamp, 17
Kinds, 14, 22
Krifka, 16, 21, 22, 23, 25
Kwakw'ala, 20
Landman, 14, 39
language universals, 2, 3, 27, 38
Lewis, 17, 29
lexical, 24
lexical categories, 3
Li, 29
Link, 14
Link, Godehard. 1983., 39

24