ON SURFACE FORM AND LOGICAL FORM*

Edward L. Keenan

UCLA and the University of Tel Aviv

The body of this paper, sections 2 and 3, presents two quite general correlations between the surface forms (SFs) of natural languages and their logical forms (LFs). Section 1 discusses briefly the nature of SFs and LFs, and section 4 assesses the significance for linguistic theory of establishing relations between SFs and LFs.

1. *The Nature of SFs and LFs.* I will assume a SF (= a surface structure) can be represented by a labelled tree (phrase marker) of the sort usual in generative grammar. Subtrees of a tree represent constituents of the SF and are themselves SFs (since they are also labelled trees). The labels on the nodes of a tree represent the grammatical categories and subcategories of expressions of the language.

It is worth emphasizing that the SFs of a language are defined by a complex set of rules which may be empirically validated (judged descriptively adequate) according as (1) the phonologically interpreted SFs they define are judged grammatical (the soundness criterion) and (2) according as all the forms speakers judge grammatical are in fact defined (generated) by the rules (the completeness criterion). In addition a set of rules may be judged explanatorily adequate according as it is useful in explaining other facts about the language, such as that children learn languages quickly on the basis of limited exposure.

Consequently there can be no simple direct argument concerning the correctness of a particular assignment of SF to a particular expression of a language. We can only assess the descriptive and explanatory adequacy of the entire set of rules, and it is this set which determines assignments of SFs in particular cases. Further, as we modify the rules to make them descriptively more adequate our assignments of SF will change (as the history of generative grammar amply demonstrates).

The nonobvious nature of structure assignment is something linguistic theories share with all scientific theories. There is for example nothing obvious, or given, or natural, about the atomic structure which physical theories assign to oxygen. Our representations here are part of a theory which assigns structures to many compounds and elements and whose correctness is determined by its predictions of the chemical and subatomic properties of these elements and compounds. Further, our assignment of structure in particular cases has undergone massive changes over time as our understanding of the chemical and subatomic properties has grown.

What holds of SFs and atomic forms also holds of LFs. Thus given a set of expressions of a language (e.g. English, elementary arithmetic, etc.)

a logical theory defines a set of LFs (logical structures) whose descriptive adequacy is determined by the correct predictions they make concerning the logical properties of the expressions of the language. For example it is a logical property of (1) below that it entails (logically implies) (3), whereas (2) does not have this property.

(1) Both John and Mary can swim

(2) Either John or Mary can swim

(3) Mary can swim

(To say that a sentence A entails a sentence B is just to say, informally, that in any state of affairs in which A is true, B is true. That is, whenever the world is the way A says it is, it is also the way B says it is.)

Furthermore, logical theories for a natural language can also be assessed as more or less explanatorily adequate according as the set of LFs the theory defines is useful in the explanation of other properties of the language. For example, in *Logical Types for Natural Language* (Keenan and Faltz 1978, henceforth *Logical Types*) it is argued on several different grounds that the LFs that theory provides is explanatorily more adequate (in addition to being descriptively more adequate) than those provided in 'The proper treatment of quantification in ordinary English' (Montague 1973). We argue for example that our LFs make fewer demands on the logical competence of speakers (and hence learners) of the language.

As with SFs then there is nothing obvious or given about the assignment of LF to a particular expression of a language. The assignment depends on the descriptive and explanatory adequacy of the logical theory which defines the set of LFs. And as with SFs, as we modify our logical theories for a natural language in the direction of increasing descriptive adequacy the assignment of LFs to expressions of the language will change. For example, a purely sentential logic for English is not rich enough to show that the a-sentences below entail the corresponding b-sentences:

(4) a. All men are mortal and Socrates is a man
    b. Socrates is mortal

(5) a. All horses are animals
    b. All heads of horses are heads of animals

First order logics, which distinguish a predicate argument structure in simple formulas and which allow quantification over arguments can represent the entailment relations given above. And recent work in natural logic, such as Montague Grammar generally (cf. Partee 1976) and *Logical Types* distinguishes more structure than the classically-given first order predicate calculus (CL) and is descriptively more adequate for English than is CL. (Richness of logical structure does not necessarily correlate with being a higher than first order logic. See section 4 for discussion.)

In addition to the theoretical and methodological properties which

scientific theories in general share, syntactic theories and logical theories have some additional properties in common. Namely the structures which both define are representable by labelled trees. That is, they are linguistic objects, having a constituent structure, etc. It thus makes sense to ask to what extent the SF a syntactic theory assigns to an English expression corresponds or 'looks like' the LFs assigned to that expression by a logical theory. Is there for example a one-one correspondence between the constituents of the SF and those of its LFs? Do constituents of different grammatical categories correspond to logical constituents of different logical categories? And so on.

Answers to these questions depend of course on the exact nature of the SFs and the LFs the two theories define. Consider for example the English SF in (6).

(6)
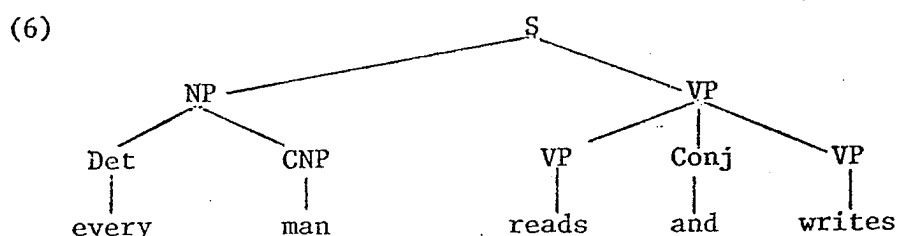
The example is for illustrative purposes and does not for example represent the nonsegmented Auxiliary or the verb agreements. Further we do not insist on the label CNP, common noun phrase, but we do insist that *every student* and *student* in (6) be assigned different category labels, a distinction recognized both by traditional grammar and X-bar notation. We shall henceforth for clarity refer to expressions like *student, fat student,* etc., as CNPs and expressions like *every student, a student, John, he,* etc. as DNP's, determined noun phrases.

A plausible LF for (6) in CL would be:

(7)
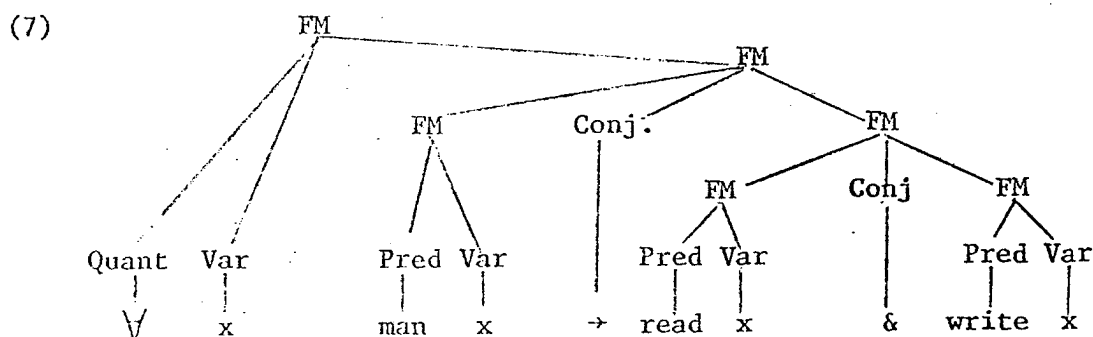
Clearly the correspondence between (7) and (6) is none too good. (7) for example contains five properly embedded formulas (FMs), the category which corresponds to Sentence, whereas (6) has no properly embedded sentences. Further the category distinction in English between CNP and VP is lost in CL, both being represented as one-place Predicate symbols.

By contrast the LF for (6) given in *Logical Types* is:

(8)

```
                              FM
               ┌───────────────┴──────────────┐
             DNP                              VP
           ┌───┴───┐              ┌────────────┼────────────┐
         Det      CNP            VP                         VP
          │        │             │            │             │
        every     man          read          and          write
```
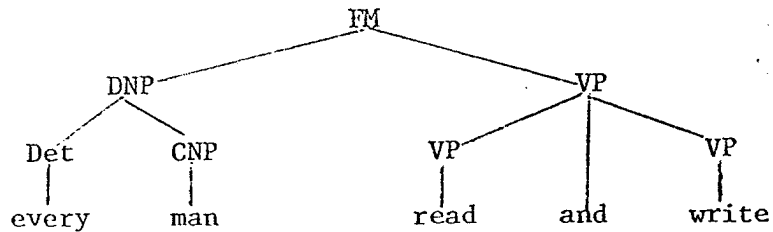
Obviously the correspondence between constituent structures is close to
one-one here, and all the category distinctions in English are preserved
in the LF. *Logical types* does not represent Auxiliaries or verb agreements
however and further the trees are unoriented. That is, left-right order is
not represented in (8). (8) is in fact a set of two expressions, the DNP
every man and the VP *read and write*. The DNP is itself a set of two
expressions whose members are *every* and *man*, and the VP is a set of three
expressions: *read, and,* and *write*.

On the other hand, (9) below would be, if I am not mistaken, a LF
assigned to (6) in Montague Grammar. (I have used where possible 'standard'
names for categories rather than those actually used in MG.)

(9)
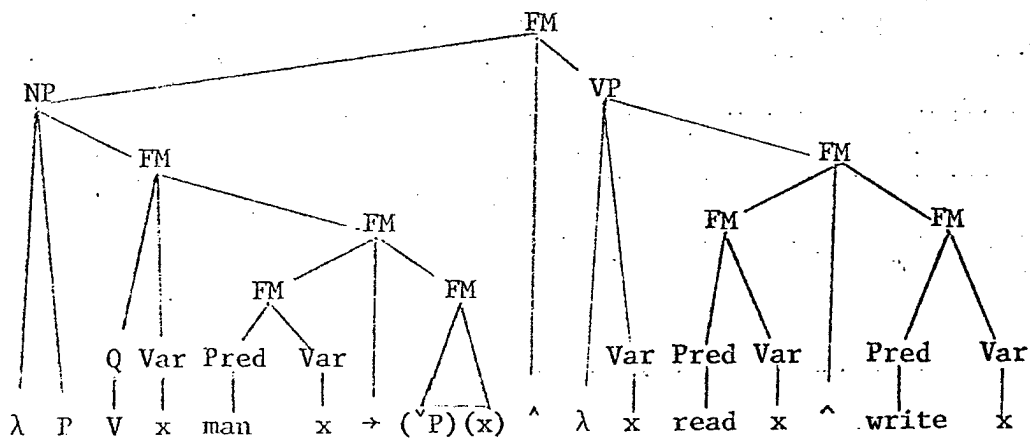
Clearly the correspondence here is less good than in CL. Not only do we
lose the distinction between CNP and verb but (9) presents seven properly
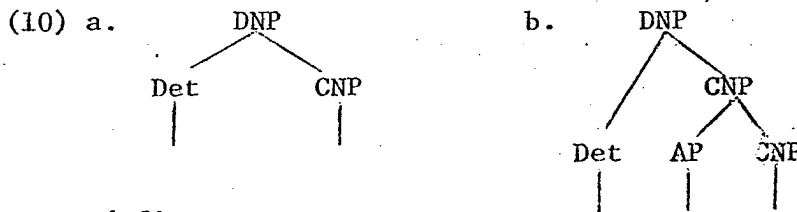embedded formulas.

Given a set of SFs and a set of LFs for a natural language we may now
say that a property P defined on the SFs <u>corresponds</u> to a property P'
defined on the LFs just in case a SF has P iff some (all) of its LFs have
P'. And in general SFs will correspond to LFs to the extent that their
properties correspond. Further, this notion of correspondence can be
naturally generalized in several ways. We may replace P and P' by n-place
relations R and R' and we may allow the arguments of R and R' to be not
simply SFs and LFs themselves but sets of SFs and LFs.

To support our claim that in general there is a significant correspon-
dence between the SFs for a natural language and their LFs we shall consider
here two relations defined on SFs, the <u>agreement</u> relation and the <u>left-
right order</u> relation and argue that in a large class of syntactically
specifiable contexts these relations do correspond to relations defined
on LFs.

2. *The Agreement Relation.* We observe across a wide range of languages that certain constituents of SFs 'agree' with certain others. For example in a SF for a sentence consisting of a VP and a DNP (regardless of their relative order) the VP may agree with the DNP. That is, a given expression of the VP will vary in form according as expressions for the DNP are drawn from different gender classes and have different numbers. In the same sense we find that transitive verbs (TVs) and in fact transitive verb phrases (TVPs) sometimes agree with their DNP objects (the DNP they combine with in whatever way to form a VP), though this agreement is less common than VP agreement with subjects. Another common case is that of adjective, and more generally adjective phrase (AP) agreement with their CNP 'heads'.

But we don't expect that just any constituent of a SF may agree with any other. It is then an interesting problem in Universal Grammar to specify the pairs of SFs <A,B> such that A may agree with B in a syntactic structure E. (By may agree I mean that there are possible human languages in which they do agree.)

To define the problem somewhat more precisely, let us consider here a syntactic structure (and so in particular a SF) to be a labelled tree of the usual sort less the terminal symbols, that is, less the actual expressions of the language. We may then say of two expressions in a language that they have, or express, a same SF. Arguably for example both *some tall man* and *every woman* express (10a) below, though only *some tall man* expresses (10b).

(10) a.
```
        DNP
       /   \
     Det    CNP
      |      |
```

b.
```
        DNP
       /   \
      /     CNP
     /     /   \
   Det    AP   CNP
    |      |     |
```

We may now define:

(11) Given A and B distinct constituents of a syntactic structure E in a language L, A agrees with B iff the form of expressions of A varies with the choice of expressions of B.

(To say that A and B are distinct is to say that neither is a subconstituent of the other.) And our problem may now be stated thus:

(12) For which choices of E, A, and B as above is it the case that there exist languages in which A agrees with B?

A reasonable starting point is the common cases of agreement cited above: VPs may agree with their DNP subjects; TVPs with their DNP objects; and APs with their CNP heads. Do these cases have anything in common? I think so. Namely, there is a sense in which how we semantically interpret the item which agrees varies with how we interpret the item it agrees with. This claim is not obvious and will be justified below. Specifically, let us consider the Meaning-Form Dependency Principle as a first attempt to characterize the class of possible agreement phenomena.

THE MEANING-FORM DEPENDENCY PRINCIPLE (MFDP)

> Given A and B distinct constituents of a syntactic structure E,
> A may agree with B iff the semantic interpretation of expressions
> of A varies with the semantic interpretation of expressions of B
> in the interpretation of E.

Loosely, the form of A's varies with that of B's just in case the interpretation of A's varies with that of B's. Which is to say that a surface relation holds of A,B pairs just in case a certain semantic relation holds between their interpretations. And this is just the kind of correspondence we are looking for, although the MFDP does not specifically mention a relation defined on LFs to which the agreement relation on SFs corresponds. We will state such a relation shortly, but as it is the MFDP which gives the intuition behind the principle, let us first see what justification there is for it.

Consider first the case Adjective Phrases. If the principle is correct it should be the case that how we interpret modifying adjectives varies with the interpretation of the CNP they modify. And this is in fact the case. Consider for example an ordinary adjective like *flat*. In expressions like *a flat road, a flat table*, etc. *flat* means something like 'without bumps or depressions'. And this in general seems to be the interpretation of *flat* when the CNP it modifies refers to a class of objects with extended surfaces. But if the CNP refers to a normally tasty liquid, as in *flat beer/champagne* then *flat* means something like 'lacking its normal taste, having lost its "zip"'. And if the CNP refers to a class of objects which, in their usable state, are inflated with a gas of some sort, as in *a flat tire/balloon* then *flat* means something like 'deflated'. And if the CNP refers singing or voice quality, as in *a flat voice, flat sopranos*, then *flat* means something like 'too low in pitch'. In short, our interpretation of *flat* is adjusted to, or varies with, the kind of thing we are predicating it of.

And this is commonly the case for common adjectives. Thus strong tables or fortresses are ones which can <u>withstand</u> much force, whereas strong wrestlers or weightlifters are ones that can <u>exert</u> much force; and strong tea or tobacco are substances which <u>produce a marked effect</u> on those who ingest them. Similarly *solid* buildings or tables are basically strong ones; but solid bars or cubes are ones which are not hollow, solid gold or silver are substances of a high degree of chemical purity, and solid lines or rows of buildings are ones without a break.

We may infer then that how we interpret modifying adjectives, and thus APs in general, varies according to the meaning of the CNP they modify.

Analogous claims hold for basic transitive verbs. Consider for example *cut*. If its object is (roughly) an animal or a largish body part of an animal, *cut* means something like 'to make an incision in the surface of'. Note that the integrity of the object cut is understood to be preserved. A finger which has been cut is still a finger, and can be cut again. Thus *cut* when applied to animals or body parts does not imply cutting all the way through. Nor does it imply intentionality or a purpose on the part of the

agent. I may cut John either accidentally or on purpose, and in each case the adverbial phrases *accidentally, on purpose* adds information not present in the meaning of *cut* itself. On the other hand, if the object of *cut* is a prepared foodstuff, as in *to cut a cake, the roast, cut* is now interpreted as something like 'to divide into portions for the purpose of serving'. So the action *cut* specifies in these cases is an intentional or purposive one. Further, the integrity of the object cut is seriously affected and normally destroyed as we do now cut all the way through. On the other hand, if the object of *cut* refers to the kind of object consisting of many uniformly elongated objects, as in *to cut the lawn, your hair, one's fingernails, cut* is interpreted to mean something like 'trim'. In distinction to the first sense of *cut* above (cut a person), we do cut all the way through in that parts of the whole are cut off. And in distinction to the second sense above (cut the cake) we do preserve the integrity of the whole. Once we have cut the lawn there is still a lawn there. And further, some notion of purpose is implied. We cut the lawn or our hair to make them more regular, pretty, or something like that.

Further, if the object of *cut* is a concrete mass noun, as in *to cut alcohol* or *cut heroin, cut* means something like 'diminish the potency of by admixing a physically comparable substance'. Thus we may cut whiskey with water, marijuana with tea, etc. No notion of making an incision is present in any sense in this interpretation of *cut*. And if the object of *cut* is an abstract one of certain sorts, as in *to cut working hours/production quotas/prices, cut* means something like 'to decrease the value of along a numerically continuous dimension (the dimension in question of course is one appropriate to the nature of the object cut). And matters are still more varied. Thus *to cut a class or a meeting* means to not attend when supposed to; *to cut a path* (through a field) or *a tunnel* (through a mountain) means to create a path or a tunnel by cutting something else and removing it; *to cut a film* or *a dissertation proposal* means to eliminate parts of it (destroying integrity) and put the result back together again (restoring integrity) yielding a shorter and presumably better work (notion of purpose involved).

In short then, the action we interpret *cut* as specifying is one appropriate to the object affected. And what holds of *cut* here holds as well of other common transitive verbs. To *drive cars or buses* is a different action from driving cattle or sheep; and both actions are different from driving a motor or a generator. *To drill* means something different in *to drill recruits* than in *to drill metal*, etc. Thus transitive verbs, and so transitive verb phrases in general, vary in their interpretation with the interpretation of the object affected.

Finally, intransitive verbs, and so VPs in general, behave similarly. Thus to say of an animal that it *is running* is to say both that the animal is moving its parts (internal movement) and that it is changing location however momentarily relative to other things (external movement). But to say that a watch or a car motor is running is at best only to make a claim concerning internal movement; no external movement is implied. If you are running with your watch it does not follow that your watch is running. And to say that your nose or the faucet is running implies neither internal nor external movement on the part of the nose or the faucet. Rather it

implies that liquid, of a sort appropriate to the referent of the subject phrase, is moving out of that referent. And to say of a play or an exhibition that it is still running is to say that it is still available for public enjoyment, with no implication of physical movement at all. In short, how we interpret *run* in a context 'DNP is still running' depends on what sort of object the DNP refers to.

Note clearly that the dependency here is clearly semantic. If we say *it is still running* how we interpret the verb depends on what *it* refers to. And if I ask you to *cut them* the action I'm asking you to perform is different according as *them* refers to production quotas or salamis. And the interpretation of flat in *a flat one* or *flat ones* depends on the class of item referred to by *one* or *ones*.

It does appear then that the MFDP holds for the common cases of agreement considered. But many questions concerning this dependency still need to be answered. At what level of linguistic description should this meaning dependency be represented? Semantic? Pragmatic? (We have remained non-committal on this point.) Does the same sort of form-meaning dependency obtain for other cases of agreement phenomena? And is there any correlation between what varies in meaning with what and the logical structures assigned to agreement pairs?

Let us consider the last question first. Within Classical Logic there seems no complete description of the agreement pairs (VP and NP, TVP and NP, AP and CNP). Simple VPs correspond to one-place predicate symbols, and a limited class of DNPs, such as proper nouns, corresponds approximately to individual constants (0-place function symbols), so we may consider that simple cases of [$_S$DNP, VP] structures correspond to atomic formulas, in which, semantically, the VP is interpreted as a function taking the DNP interpretation as an argument. But the correspondence breaks down when more complex DNPs in English are considered. Thus *every man is running* does not correspond to a formula with a function-argument structure even grossly isomorphic to that of *John is running*. The two cases may be compared as follows:

(13) a. John is running
     b. R(j)

(14) a. Every man is running
     b. $(\forall x)(M(x) \rightarrow R(x))$

In fact, nothing in (14) corresponds to (interprets) the English DNP *every man* at all.

Transitive verbs correspond somewhat to two-place predicate symbols in CL, but a VP in English consisting of a transitive verb plus object does not correspond to any logical category in CL. On the other hand, it is easy to reformulate CL without changing the entailments of any of the formulas in such a way that transitive verbs do correspond to something which combines with a simple DNP to form a VP. A classical two-place predicate is semantically interpreted as a function U^2, the set of ordered pairs on U, the universe of discourse (i.e. the entities which exist in

that state of affairs) into the set of truth values, represented as 0 (false) and 1 (true). Given such a function f, there is a unique other function g which maps members of U onto VP meanings, that is functions from U into {0,1} such that for all x,y in U, $(g(x))(y) = f(y,x)$. So CL could be trivially reformulated so that, for a very simple class of DNPs, transitive verbs in English would correspond to functions taking DNP meanings as arguments, and thus, for this limited class of DNPs, there would be a similarity between VP and TVP representations in CL. Namely, both would be functions taking DNP meanings as arguments.

But when the AP,CNP pairs are considered CL has nothing to offer us. Expressions like *flat road* are not directly assigned any logical category. Nor is there anything in CL equivalent to a modifying adjective. The best translation in CL of (15) below would be (16).

(15) Every happy man is laughing

(16) $(\forall x)(man(x)\ \&\ happy(x) \rightarrow laugh(x))$

So the correspondents of 'happy' and 'man' do not form a logical constituent, and are in fact of the same logical category (one-place predicate), and that category is the same as that for simple VPs *(laugh)*.

Montague Grammar might appear to offer a more promising set of logical structures since at least the formal English (which is 'translated' into the formulas of an intensional logic and only then interpreted) does have categories of AP and CNP. However, even ignoring the fact that the distinction between CNP and VP does not survive the translation into the set of logical structures, I can still find no basis for a similarity between the three agreement pairs we have considered. It is the case that APs and CNPs form logical constituents, and the APs are interpreted as functions taking CNP meanings as arguments. And it is also the case that TVPs and DNPs form a constituent in the logic and TVPs are interpreted as functions taking DNP meanings as arguments. So for these cases we can say that English item which shows agreement is in each case interpreted as a function taking as argument the (interpretation of) the item it agrees with. But the similarity breaks down when the NP,VP pairs are considered. For here, while such pairs are a logical constituent (formulas), it is not the case that the agreeing item, the VP, is interpreted as a function taking as argument the item it agrees with. Rather the DNP is interpreted as a function taking the VP meanings as arguments.

On the other hand, in *Logical Types* there is a uniform function-argument assignment. VPs are interpreted as functions on their DNP Subjects, TVPs as functions on their DNP objects, and APs as functions on their CNP arguments. Further, the function-argument assignment holds regardless of whether the NPs in question are simple ones like proper nouns, or complex ones like *every flat road*. Thus the LF assigned on that theory to *every handsome man loves a beautiful woman* is:

(17)

```
                              FM
                 /                      \
              DNP                        VP
           /       \                  /      \
        Det         CNP            TVP        DNP
         |         /    \           |        /     \
         |       AP      CNP        |      Det      CNP
         |        |       |         |       |      /    \
         |        |       |         |       |    AP      CNP
         |        |       |         |       |     |       |
       every  handsome   man      love      a  beautiful woman
```
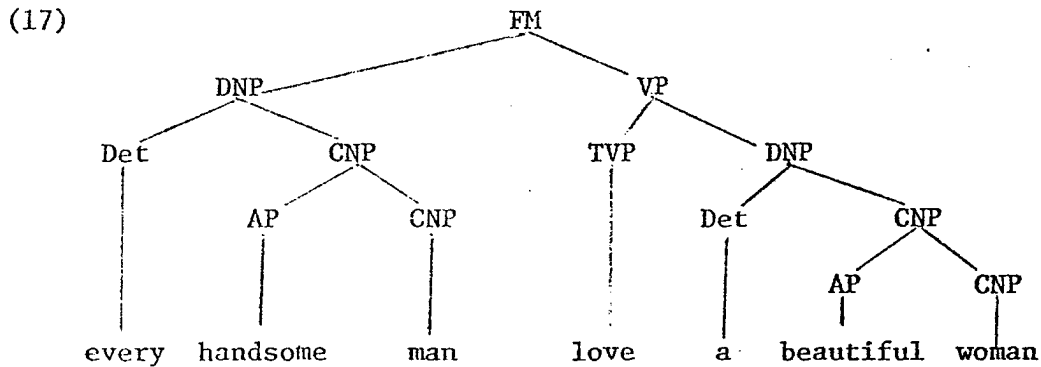
We should stress here that (17) is itself a LF in *Logical Types*; that is, it is directly interpreted, not mapped onto a formula in some intensional logic, that is, a different structure, and then interpreted.

Assuming the LFs defined in *Logical Types* we may directly characterize the correspondence between agreement pairs and LFs by:

THE FUNCTIONAL DEPENDENCY PRINCIPLE (FDP)

Given A and B distinct constituents of a SF E, A may agree with B iff in the LF of expressions of E, the LFs of expressions of A are interpreted as <u>functions</u> taking the interpretations of expressions of B as arguments.

In such cases we shall refer to expressions of A as <u>functional expressions</u> and expressions of B as <u>argument expressions</u>.

Further, as *Logical Types* defines a rather large number of function-argument structures involving DNPs and CNPs, we have at hand a ready set of predictions concerning possible agreement phenomena involving these categories. Let us note the following assignments given in *Logical Types*:

(18) Functions taking DNP arguments:
    a. VPs
    b. TVPs
    c. Possessive Phrases (e.g. *the father of*, *'s father*)
    d. Prepositions

(19) Functions taking CNP arguments
    a. APs
    b. Relative Clauses
    c. Articles (*the, a*)
    d. Quantifiers (*every, no*)
    e. Numerals (*one, two*, etc.)

Considering the second category first, the FDP predicts the existence of languages in which APs agree with CNPs, which we have already seen to be correct. Second, it predicts agreement between Relative Clauses and their CNP heads. This is also correct, for recall the definition of agreement. A agrees with B iff the form of expressions of A varies with the choice of expressions of B. Now a great many languages form Relatives by presenting

in the position relativized a personal pronoun agreeing in noun class with the CNP head of the Relative, as illustrated below from Hebrew:

(20) a. *ha- ish* [*she- natati lo et ha- sefer*]
             RC
    the-man that-I+gave to-him DO the-book
    'the man that I gave the book to'
  b. *ha- isha* [*she- natati la et ha- sefer*]
             RC
    the-woman that I+gave to-her DO the book

Clearly the forms of the Relative Clauses vary (one contains *lo* where the other has *la*) according to the gender class of the head CNP. Similarly languages which use special forms of pronouns in Relatives, so called relative pronouns, often present cases where the relative pronoun agrees in gender and number with the head CNP, as thus the form of the entire RC varies with certain choices of expressions for the CNP. So the prediction is verified. Note however, that the FDP does not predict in this case which part of the Relative Clause will vary in form but only that the Relative Clause as a whole will vary.

Third, the FDP predicts that articles may agree with the CNPs they are functions on, and such cases (e.g. Romance) are well known, though perhaps less widespread than the accessibility of confirming cases would indicate.

Fourth, the FDP predicts that quantifer words like *every* and *no* will agree with the CNPs they combine with, and again confirming instances are easy to come by:
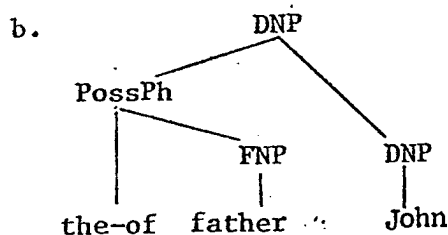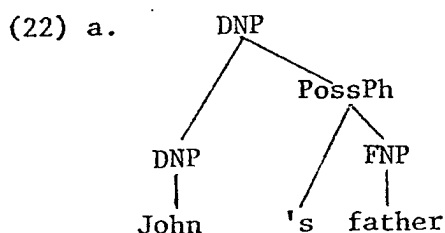
(21) a. *tous les hommes*   b. *toutes les femmes*
     all   the men   all   the women

And finally, the FDP predicts that Numerals may agree with their CNPs. Again this is correct; though it is very common that the internal structure of numerical DNPs is complex and nonuniform from one number to another. Nonetheless Russian provides a confirming instance. The numeral 'one' has four forms, one plural form (*odnix*, used before inherently plural CNPs of the sort *pants*, *scissors*, etc.) and three singular forms, one for each of the three genders of CNPs. The amount of agreement decreases rapidly as we consider the higher numbers, but still 'two' in the singular makes at least two gender distinctions.

Recalling now that the motivation for the function-argument assignment given in *Logical Types* (or any logic) is semantic—that is, given that function-argument assignment we get many right and few wrong predictions of entailments, the fact that all the predictions in (19) are correct is rather striking confirmation of the FDP.

Consider now (18). The FDP predicts, correctly as we have already seen, that VPs may agree with their Subjects, and TVPs with their Objects. Further, within the logic, heads of possessive constructions are assigned a logical category and are semantically interpreted as functions on the Possessor NP. E.g. *John's father* and *the father of John* would be assigned

the following LFs:

(22) a.
```
              DNP
             /  \
            /   PossPh
           /    /  \
        DNP    /   FNP
         |    /     |
       John  's   father
```

b.
```
                  DNP
                 /   \
            PossPh    \
            /  \       \
           /   FNP     DNP
          |     |       |
       the-of father  John
```

(FNP abbreviates Function Noun Phrase, a primitive category in the system.)

And the FDP predicts, correctly, that heads of possessive constructions, Possessive Phrases in *Logical Types*, may agree with their possessors. This in fact is massively common in Amerindian languages, and is evidenced in many others as well, e.g. Turkish, Arosi (Melanesian), Daga (New Guinea), and Hebrew. (23) below from Hebrew is illustrative:

(23) a. *beit-o   shel Dan*   b. *beit-a   shel Miriam*
        house-his of Dan   house-her of Mary
        'Dan's house'   'Mary's house'

It is perhaps fair to mention here that while the form of possessive agreement cited above is very widespread—my impression is that it is as widespread as is adjective agreement with CNPs and TVP agreement with their NP Objects—there are some limited cases in which the agreement goes the other way. One obvious case are the pronominal possessives in Romance. Thus we have *son frère* 'his/her brother' and *sa soeur* 'his/her sister'. The possessives here behave in fact like adjectives as regards agreement, and are traditionally called possessive adjectives. The paradigm however is limited to the pronominal possessive forms. The productive possessive construction shows no such agreement, e.g. *le frère de Jean* 'the brother of John' and *le frère de Marie* 'the brother of Mary'. A somewhat larger class of 'wrong' agreements can be found in certain Slavic languages where a DNP with a possessor reading may, in certain cases, be constructed with an adjectival ending and thus agree with the head CNP as an adjective would. Comrie and Thompson (1978) cite the following from Czech:

(24) a. *vědcova   kniha*   b. *?kniha vědce*
        scientist-f.sg. book   book scientist-gen
        'the scientist's book'   'the scientist's book'

Note that (24a) clearly has a possessive meaning, not something like 'a scientific book' (which would be *vědecká kniha*). And contra the FDP the logical possessor shows agreement with the CNP head. Again however the productivity of this paradigm is limited, in fact listable (though the list would be much longer than for the Romance cases cited above). Thus only definite, singular, unmodified possessors can be construed as adjectival. To say *the old scientist's book*, *the (many) scientists' books*, or *a scientist's book* we could only use the 'proper' possessive construction illustrated in (24b). Bob Rothstein (personal communication) points out however that at least certain proper noun possessives, as *Newton's theories*, may also be constructed with the possessor as an adjective.

It would seem then that to maintain the FDP as stated we should weaken it to allow for listable exceptions. And this is not unreasonable. If we establish a correlation between the most productively generated SFs and their LFs we have surely established a correlation between SF and LF in general. Exceptions to the general pattern will have to be learned as special cases. Indeed for other types of correlation between SF and LF, such as some version of the Fregean Principle (meaning of the whole is a function of the meaning of the parts) various types of fixed expressions—proverbs, idioms, etc.—will clearly be listable exceptions.

Finally, FDP predicts that prepositions may agree with the NPs they govern, although the LFs in *Logical Types* are less clear here, as Prepositions have multiple categories.[1] Thus in addition to combining with DNPs to form various classes of modifiers they may also combine directly with VPs and TVPs to extend the class of DNP arguments they take. Thus in *Logical Types* (*sleep, in*) is a derived TVP requiring a DNP object to make a VP again (so [(*sleep, in*), (*the bed*)] is a derived VP). Further, cases of Preposition agreement with DNPs are harder to find across languages, but some cases do exist. (25) below from Arosi (Melanesian; see Capell 1971 for supporting details) is one example:

> (25) *'ini- a mada*
>   with-it club
>   'with a club (he hit me)'

Worth noting here is that not all prepositions show agreement with the NP they govern, and those that do, show an agreement paradigm similar (perhaps identical) with that that transitive verbs show with their Objects. However, given the similarity in function-argument structure between transitive verbs and their Objects and Prepositions and their Objects it is perhaps not surprising that transitive verbs can historically become reduced to function as dependent forms, i.e. Prepositions. Note as well that another very common source for prepositions is possessive constructions. Thus for example *behind it* will in many languages be something like *at its back*. If the language were such as to have agreement between possessive heads and Possessors we could expect that prepositions so derived historically might show such agreement as well. Something like this may be the explanation for the postposition agreement in Daga (New Guinea; see Murane 1974 for details). In any event there are several clear cases of postpositions agreeing with their DNP Objects, and the agreement suffixes are the same as the intimate possessive agreements with possessors:

> (26) *ge orup ame ena- m ak*
>   you fellow that with-their go+2sg+imp
>   '(You) Go with those fellows'

Thus for all the function-argument structures in *Logical Types* which concern either DNPs or CNPs, the Functional Dependency Principle correctly predicts the possibility of agreement. And as these nine cases represent a reasonably large sample of the possible agreement phenomena known to me, I feel that even if the FDP were restricted to hold for just these cases it would still represent a very significant correlation between SF and LF and thus support the general claim that there exist correspondences between

SF and LF. On the other hand, a large number of questions concerning the FDP remain. In section 2.1 below we raise some of these questions.

2.1 *Refining the FDP.*

2.1.1 We have replaced the original intuition expressed in the Meaning-Form Dependency Principle by one, the FDP, stated in formal logical terms. It is reasonable to ask how well the original intuition behind the MDFP carries over to the larger range of cases. The results here are mixed and interesting, and suggest some additional correlations between SF and LF.

Consider first the Possessive Phrases. Do the heads of such constructions vary in their interpretation with the reference of the Possessor in anything like the same way in which say *flat* varies with the interpretation of the CNP it modifies or *cut* varies with the semantic nature of its Object? Here there does seem to be some variation, though it is not clear that it is of exactly the same sort as the very idiosyncratic variation exhibited by *flat, cut,* and *run*. Thus consider a PossPh like *the middle of.* If its argument, the Possessor, denotes an activity, such as the performance of a play or a radio program, then *the middle of* picks out that part of the performance which occurs at a moment, or period, of time midway between the beginning and the end of the activity. Analogous claims hold for arguments that denote specified periods of time, such as *day, night*, etc. But if the argument refers to a concrete physical object *the middle of* has different senses. Thus the middle of the Earth is a physical point, or region, more or less equidistant from all the points on the surface of the Earth. So *the middle of* may have a temporal or physical interpretation according to the nature of the Possessor—a fact which is well known. But this variation is not wholly similar to that evidence for *flat, out*, etc. That is, the same variation shows up for a few other such expressions, e.g. *the end of, the length of,* whereas I don't expect (but haven't in fact checked) that any other adjective besides *flat* would have exactly the four senses attributed to it earlier. On the other hand, *the middle of* does show further different senses which do seem more idiosyncratic when applied to different sorts of physical objects. Thus if the Possessor refers to something conceived of in two dimensions, *the middle of* may pick out a *point* at the geometrical center, as in *the middle of the stage* or *the middle of a sheet of paper*. But if the two dimensional object is notably elongated, *the middle of* picks out a *line* more or less equidistant from the edges, as in *the middle of the road*. Further, if the object is conceived of in three dimensions, then differences in the interpretation of *the middle of* show up according as the object is solid or one habitually entered by human beings. Thus *the middle of the Earth (the Sun) consists of molten iron (Helium atoms)* means that the space at the geometrical center is filled with molten iron, etc. But *John is sitting in the middle of the room* does not suggest that John is suspended in midair. Rather the middle of the room defines a region of space with reference to the center of the (two dimensional) floor.

In addition, many overtly possessive constructions lend themselves easily to 'metaphorical' interpretations of the head (or PossPh). Thus one might argue that *the head of* in *the head of the animal* is interpreted literally, and picks out a certain body part of the animal. But *the head*

of the table or *the head of the department* force *the head of* to be understood in a nonliteral (metaphorical) way. This sort of metaphor is what Reinhart (1976) calls nonpoetic metaphor and is characterized, in our terms, by a reinterpretation of the function expression given a 'nonordinary' argument expression (the Possessor). But whether we take one sense as basic and the others as metaphorical or not, it is still the case that it is the function expression which is subject to various interpretations depending on the (literal) meaning of the argument expression.

In addition we might note that nominalizations of various other sorts of functional expressions (adjectives, VPs, etc.) commonly take the form of possessive constructions, and some, but not all, of the sense variation in the original may be preserved under the nominalization. Thus *the strength of* has a similar range of senses to *strong* (person, fortress, tobacco, etc.). On the other hand, *the solidity of* loses some of its senses. Thus while *the solidity of the chair* presumably does refer to its capacity for withstanding force, *the solidity of the bar* or *the cube* is unclear in meaning, referring most likely in my opinion not to its quality of being not hollow but to its capacity of withstanding force. And the solidity of the gold or silver cannot be used to refer to its chemical purity but must refer to its capacity to withstand force, if indeed it refers to anything at all. Nor can *the solidity of a line* or a row of buildings naturally refer to their quality of being 'perceptually unbroken'. Note also that *solid* loses its chemical purity sense when used predicatively. *This gold is solid* does not mean it is chemically pure. These observations suggest:

THE FUNCTIONAL DEPTH PRINCIPLES

(i) The meaning range of functional expressions decreases in proportion to its internal function-argument complexity, and

(ii) for f a functional expression and e an argument expression, the meaning range of f(e) is less than or equal to that of e.

Thus (i) says basically that more derived items exhibit less meaning variation than less derived items of the same category. And (ii) says that a derived item exhibits less (or at least not more) meaning variation than the item it was derived from.

Thus (ii) would predict for example that *run slowly* should have a narrower (or at least not greater) meaning range than *run*, given that *slowly* is interpreted as a function on the interpretation of *run*. And this appears to be correct. While we can understand *John/the motor/the faucet is running slowly* using *run* in the same sense as if *slowly* were absent, we cannot say *My Fair Lady* or *the Braque exhibition is running slowly*, so *run* has lost one of its senses.

(i) would predict that VPs consisting of a TVP plus object would exhibit less meaning variation than simple lexical intransitive verbs. And this also seems correct. It is not easy to find TVP + DNP combinations in English which exhibit the gradient, idiosyncratic variation exhibited by *run*, though some variation still does exist. Thus *to cross the river here* is interpreted as an activity if the subject is animate (e.g. *the soldiers*), but it is interpreted as denoting a state or a locative relation with

various choices of inanimate subject, e.g. *the bridge, the telephone lines,* etc. The Depth Principle also suggests that a predicate adjective may have fewer senses than its modifying use since, e.g., *be solid* is functionally more complex than *solid*. And at least for *solid* this appears to be correct.

The Depth Principle very likely has something right about it then, though much more work investigating particular cases would be needed to place it on firm ground. Furthermore there are other restrictions on the meaning range of functional expressions besides those covered in the Depth Principle. Thus some lexical adjectives (intransitive verbs, etc.) simply exhibit much less meaning variation than do others. Thus the *Random House College Dictionary* (1973) lists only one sense for *eleemosynary*, two for *perspicacious* (one archaic), and three for *melancholy*, whereas it gives 23 for *flat*, 23 for *strong*, and 26 for *solid*.

A kind of limit case here may be *logical constants*, items like *every* and *not* whose interpretations do not vary from state of affairs to state of affairs. And as *Articles*, *Quantifiers*, and *Numerals* are logical constants in this sense we would perhaps not expect to exhibit meaning variation. And that appears correct. We shall not here present and substantiate a principle which would predict that however. Prepositions, few in number in any language compared to for example CNPs and VPs, seem to lie somewhere between ordinary 'content' items and logical constants. And it is unclear to me whether we want to say that *with* in *to fill the tub with Harry/with water/with joy* exhibits meaning variation or is simply multiply ambiguous.

So more research is needed here. But it does seem to me likely that some version of the MFDP, suitably restricted to allow for inherent restrictions on the meaning ranges of specific items, can be made to work.

2.1.2 It is reasonable to query whether functional expressions taking items other than DNPs and CNPs as arguments also show agreement. If not, the FDP should be restricted to only cover nominal agreement. But our linguistic knowledge is limited here, so we shall not at the moment impose such a restriction. Further, we can find cases where adverbs and PPs inflect for categories that the VPs (and TVPs) they are functions on also inflect for. In Malagasy for example locative complements of verbs are marked for tense, and their marking must be the same as that of the verb:

(27) a.

$$n- \quad andeha \begin{Bmatrix} t- \\ *\emptyset- \\ *ho- \end{Bmatrix} \quad any \quad Antsirabe \quad Rabe$$

past-go $\begin{Bmatrix} \text{past} \\ \text{pres} \\ \text{fut} \end{Bmatrix}$ there Antsirabe Rabe

'Rabe went to Antsirabe'

And in Avar (NE Caucasian; see Anderson 1977) certain adverbial constructions agree with the subject, as do the verbs:

(28) a. *emen   roq'o-ve   v-us:ana*
       father home- m   m-returned
       'Father returned home'

b. *ebel   roq'o-je j-us:anax*
   mother home- f f-returned
   'Mother returned home'

But in the Avar case it is 'surely' more natural to consider that it is the entire VP *returned home* which is agreeing in gender with the subject. The apparent agreement between the VP and adverb is an artefact of how VP agreement is internally realized and not a case of the form of the adverb varying with the choice of verb, as would be required by the FDP.

The Malagasy case is perhaps less easily explained away. Given sentences like *John works and has always worked very hard*, where plausibly VPs with different tense marking occur within the scope of the adverb, it is reasonable to think that tense marking is a property of verbs, and so the adverbs or PPs which must have the same tense marking may well be instances of agreement as per the FDP. But it is also not implausible to think that tense marking is a property of the whole VP or even the entire sentence, in which case apparent agreement in tense internal to the VP would be a result of spreading of the marking on the VP as a whole and not a case of the form of the PP varying with the choice of verb.

A clear case of feature spreading rather than agreement is given by case marking on modifying adjectives. While it would be tempting to consider, with traditional grammar, that modifying adjectives agree in case with their CNP heads, it seems clear (see 2.1.5 below) that case marking is a process which takes the entire DNP in its domain and that in certain languages the case marking on the DNP spreads onto adjectives and determiners. So agreement in case, as opposed to agreement in gender class, is not a case of the form of one expression varying with the choice of another as the FDP requires.

The cases of feature spreading above neither support nor refute the FDP, but they do in certain cases support the <u>phrasal</u> as opposed to the merely lexical nature of agreement phenomena. Thus the FDP claims that VPs in general, not just lexical VPs (intransitive verbs) may agree with their DNP subjects. The claim is not falsified if only lexical VPs show agreement since variation in form of part of the VP is variation in form of the whole VP, but the FDP still does suggest that other parts of the VP may vary in form since it does not restrict agreement to lexical VPs.

Thus the examples of verb and adverb agreement with subjects cited for Avar above clearly illustrate full VP agreement as opposed to merely lexical VP agreement. (29) and (30) below from English also illustrate full VP agreement.

(29) a. He *is* behaving *himself*
     b. She *is* behaving *herself*
     c. They *are* behaving *themselves*

(30) a. He *has* lost *his* way
     b. She *has* lost *her* way
     c. They *have* lost *their* way

Note that the VP agreement is still evidenced in cases where the lexical verb itself does not agree, as in (31) below:

(31) a. They should/might/must/won't/behave *themselves* (*himself/
*oneself)
b. She may/must/should/have lost *her* way (*his way/*their way)
c. She tried/promised (Bill)/wanted/to lose *her* way

Similarly both copular verbs and predicate adjectives and predicate nominals show agreement with Subjects in many languages.

(32) a. John *is a* doctor
b. John and Mary *are* doctors

(33) a. Marie *est* intelligent*e*
b. Marie et Francoise *sont* intelligent*es*
c. Jean et Marie *sont* intelligent*s*

(33b,c) show as well that what the form of the VP varies as a function of is the entire DNP Subject and not just the 'head'. To know whether *intelligent* is feminine plural, (33b), or masculine plural, (33c), we must in principle check the gender of each conjunct of the Subject. The point is perhaps even more obvious in languages in which full verbs agree in gender, as the Hebrew examples below illustrate:

(34) a. *Miriam v-Ruti medabrot*
Mary and Ruth are speaking (f.pl)
b. *Miriam v-Dan medabrim*
Mary and Dan are speaking (m.pl)

Overall then the fundamentally phrasal (rather than merely lexical) nature of the agreement phenomena that the FDP predicts seems to us basically correct. On the other hand, whether the FDP makes correct predictions concerning function argument structures not involving DNPs and CNPs must await further logical and linguistic research.

2.1.3 But what about further function-argument structures which do involve DNPs and CNPs? A few plausible cases do come to mind. Thus we strongly expect that ditransitive verbs like *give* and *hand* will be treated semantically as functions taking, at some level of analysis, their Indirect Objects (IOs) as arguments. Assuming that, the FDP would predict the possibility of ditransitive verb phrase agreement with IOs, and this is in fact correct. E.g., Spanish, Basque, and Daga (New Guinea) all evidence such agreement.

What about the logical functions represented by *and, or,* and *not*? In *Logical Types* these operators do directly form derived DNPs, but they are treated as operators of a different sort from those already considered. The main difference is that they also combine directly with Formulas, VPs, TVPs, and APs to form derived members of those categories. Further, *and* and *or* are two-place functions rather than the one-place ones considered so far. Linguistically, relative to the category they combine with, a certain minimal variation is known. That is, many languages, for example,

use different morphemes for *and* according as the category of items conjoined is different. The examples below from Malagasy are illustrative, and Payne (1978) provides large numbers of other examples.

(35) a. *mihinam-bary Rabe ary (*sy) misotro taoka Rasoa*
    eat-   rice Rabe and   drink   booze Rasoa
    'Rabe is eating rice and Rasoa is drinking booze'
   b. *mihinam-bary sy   (*ary) misotro taoka Rabe*
    eat-   rice and   drink   booze Rabe
    'Rabe is eating rice and drinking booze'
   c. *mihinam-bary Rabe sy (*ary) Rasoa*
    eat-   rice Rabe and   Rasoa
    'Rabe and Rasoa are eating rice'

On the other hand, I know of no language in which the form of *and* varies with the choice of conjuncts within a category. Thus, as far as I know, if a language has an overt conjunction for DNPs at all then it uses the same form for *John and Bill* as it does for *Mary and Susan*.

2.1.4 We have considered a rather large number of function-argument structures and at least those which involve DNPs and CNPs are commonly enough expressed in SF by items which satisfy the agreement predictions of the FDP. But are there cases of agreement in SF which are not expressed by function-argument structures of the right sort? There would appear to be at least one: That of pronominal agreement with full DNP 'antecedents'. As 'bound pronouns' can occur in constituents in which their antecedents do not occur, the pronoun will not in general be represented as a function taking the interpretation of the full DNP as an argument. This at least is not the case for pronominal representation in *Logical Types*, nor is it the case in any of the commonly used versions of LFs for natural languages or formal languages. Rather pronominal co-reference i represented by the use of Variable Binding Operators (VBOs), illustrated in (36) below from Classical Logic and (37) from *Logical Types*.

(36) a. Every student laughed and cried
   b. $(\forall x)$ (student(x) → laughed(x) and cried(x))

(37) a. Every man loves his mother
   b. ((every,man),($\lambda x(x,$(love,$(x$'s mother))))))

The use of such VBOs seems 'unnatural' in the sense that they do not seem to correspond to anything in SF. Indeed we count it as an advantage of *Logical Types* that very many SFs using quantified NPs which bind the reference of more than one position can be represented without the use of VBOs. For example, (36a) above would be represented in *Logical Types* by:

(38) ((every,man)(laughed, and, cried))

However in the LF for (37a) there is no way to bind *his* without using the lambda operator. And it is clearly not the case that what corresponds to *his*, namely the bound occurrence of x, is interpreted as a function taking the interpretation of (*every,man*) as an argument. Similarly the position relativized into in the representations for Relative Clauses in *Logical Types* is represented by a bound variable.

The best tentative conclusion I can reach concerning these phenomena is that the kind of agreement we see between pronouns and full NP antecedents is different in kind from that that we have considered so far, and a logical characterization of it should proceed along lines different from those in the FDP. I would propose then that the FDP be restricted so as to require that the distinct constituents A and B of E are further required to be of distinct categories (not merely subcategories of one and the same larger category). Since pronouns and full NPs are all DNPs, the FDP would then make no prediction about their agreement possibilities. An additional principle would be required then to characterize pronominal agreement. The principle is in fact easy to state, though we shall not be concerned with its details or justification here. But roughly, if A and B are of the same major category, then A is represented in LF by a bound variable and B is what binds it (i.e., in *Logical Types*, B is the argument of the structure created by that VBO).

2.1.5 A final potential restriction on the FDP concerns case marking of DNPs. In many languages a DNP Subject may have one case marker, the absolutive, if the main verb of the VP is intransitive, and another marker, the ergative, if the main verb is transitive. So the form of the DNP would vary with that of the VP, in violation of the FDP. Similarly in many languages Experiencer 'Subjects' will be marked like IOs (that is, dative) whereas Subjects of activity VPs will be nominative (or unmarked).

Appearances then might suggest that the FDP should be restricted so as to cover just the traditional agreement cases and not, as Emmon Bach points out to me, traditional cases of government. To impose such a restriction we might for example stipulate that the variation in form mentioned in the FDP be restricted to variation with respect to inherent nominal properties such as number and gender. Even so restricted the principle would correctly account for a large number of specific correspondences between SF and LF and so still support the general conclusion that there exist correspondences between SF and LF.

I doubt however that such a restriction need be imposed. Even a brief look at case marking properties of languages shows, as I shall briefly argue below, that case marking is a sentence-level phenomenon. That is, case marking, despite the fact that it sometimes (see below) shows up morphologically bound to DNPs, is a property of sentences or more exactly a relation between a DNP and the entire sentence in which it occurs, and not a property of DNPs given as a function of the verb in the clause in which they occur.

To support this claim, I shall summarize a very large number of cases in which DNP case marking is independent of the choice of the verb. Thus for a given choice of verb a given DNP may have any of several case markers (CMs), and the conditioning factors appear to be almost any conceivable property of the sentence containing the DNP. (The latter claim is surely too strong, but a quick look at case marking systems reveals an astounding variety of such conditioning factors.)

First, it is obvious that locative cases are not predictable from the verb generally. Thus in the same way that in English we may say *John fell/*

on the sidewalk/near the station/in the bathroom, etc., so in case marking languages in the narrow sense, such as Finnish, a given verb like *fall* may take DNP complements in any of various locative cases. We may then dismiss locative cases, and indeed all the semantically rich oblique cases (e.g. instrumentals) from consideration and concentrate on case marking affecting Subjects and Direct Objects.

But even here there are a massive number of cases where the CM on a DNP may vary with the verb held constant.

First, consider the very widespread use of 'definite DO marking'. Many languages, such as Hebrew, Turkish, Persian, Spanish, Malagasy, etc., present DOs of the same verb in different cases according as the DO is ± definite and/or ± animate, the exact conditioning factors varying from language to language. (39) from Hebrew is illustrative:

(39) a. *ani ohev yeladim*
     I   like child-pl
     'I like children' (no CM on the DO)
  b. *ani ohev et ha- yeladim ha- ele*
     I   like DO def-child-pl def-those
     'I like those children'

A second, very widespread phenomenon among case marking languages concerns differential marking according as the DNP is understood to be partitively or wholly affected. Thus the differences discussed in Anderson (1971) as between *chew the meat/chew on the meat, shoot Bill/shoot at Bill*, etc., show up in the case marking in many languages. The case of Finnish is well known. DOs of many affirmative transitive verbs may be partitive in case if they are only partially affected by the action (and so the action may be understood as incomplete). If they are wholly affected, the CM may be either 'accusative' or nominative depending on other factors to be discussed shortly (see Timberlake (1975 and 1977) for detailed discussion). Edith Moravcsik (1978) points out furthermore that this distinction exists in a great many languages and includes Subjects of intransitive verbs as well. She cites, for example, (40) from Russian:

(40) a. *kash-a   ostalas'   na stole*
     kasha-nom remained   on table
     'Kasha remained on the table'
  b. *kash- i   eshcë   ostalos'*
     kasha-gen still   remained
     'Some kasha was still left'

Similarly many of Fillmore's pairs show up in the case marking of heavily case marking languages. (41) below from Hungarian is also taken from Moravcsik (1978):

(41) a. *János beültette   a   kertet   fákkal*
     John planted (def) the garden(acc) trees+with
     'John planted the garden with trees'
  b. *János fákat ültetett   a   kertbe*
     John trees planted(indef) the . garden+into
     'John planted trees in the garden'

(Note that as per the FDP, the verb in (41a) agrees with the definiteness of the DO, but does not agree with indefinite DOs as in (41b).)

Third, note that in some ergative languages intransitive subjects may sometimes take either the absolutive case or the ergative case, with a difference in meaning: the ergative indicating Subject intention or responsibility, the absolutive indicating accidental action. E. Pomo (Uto-Aztecan; Moravcsik 1978) and Bats (Caucasian; Comrie 1973) are examples. (42) below is from Bats:

(42) a. *so   wože*
        I-abs fell
        'I fell (not implied that it was my fault)'
     b. *as   wože*
        I-erg fell
        'I fell (it was my fault)'

Further, even a bare animacy distinction may occasionally trigger different CMs on transitive Subjects. (43) from Gugu-Yalanji (Australia) is cited from Gary and Keenan (1977).

(43) a. *dingkar-angka   kaya kunin*
        man-   erg+anim dog hit
        'The man hit the dog'
     b. *kalka-bu   kaya kunin*
        spear-erg+inanim dog hit
             (= instr)
        'The spear hit the dog'

The above cases are all ones in which semantic properties of the marked DNP are reflected in the CM but are not predictable from the verb itself. Some of these properties, such as definiteness, may also have a syntactic characterization, in which case it is not easy to decide whether the conditioning factor is semantic or syntactic. But other cases are more clearly syntactic. Thus, as is well known, many languages may case mark full DNPs and pronouns differently. Thus English does not mark full DNPs but does mark pronouns. And Dyirbal (Dixon 1972) and many other Australian languages may case mark full DNPs on an ergative-absolutive basis but pronouns on a nominative-accusative basis.

Furthermore, if a language distinguishes noun classes and also case marks, the form of the case marking may vary with the noun class. E.g., Russian and Latin case marking varies according as the nominal marked is masculine or feminine, singular or plural, etc.

The above collection of facts shows that case marking is sensitive to a large number of syntactic and semantic properties of the DNP affected by the marking, where these properties are not predictable from the verb. Further case marking will also be sensitive to whether the DNP in question forms a logical constituent with a VP or a TVP, and will need to have information regarding the semantic and/or syntactic subclass of the verb. Yet even with all this information the CM is not completely predictable. It still might, for example, mark the DNP as partitive or not, as in the

Russian cases above, or it might mark it as ergative or absolutive, as in the Bats cases above. In either case the CM itself brings in new information (partial affectedness, purposive activity, etc.). All this rather suggests the in fact quite intuitive view that case marking is both syntactically and logically a relation between a DNP and a VP or TVP. And if case marking is treated as a two-place function taking DNPs and verbs of various sorts as arguments then the FDP simply has no prediction to make, except the correct one that the form of the CM may vary with the nature of the arguments (stative vs. activity VP, definite vs. indefinite DNP, etc.).

Even this rather appealing picture however is very badly oversimplified. Evidence from very many languages shows that case marking must have much information available concerning the structure of the entire sentence in which the marked DNP occurs.

Thus, for example, several different cases are available where the CM on a given DNP varies with properties of other DNPs in the clause. Recall the Finnish cases mentioned above. The Object of many affirmative transitive verbs if not semantically partitive, and so marked by the partitive case, may still be either nominative or 'accusative'. The conditioning factor here is whether the transitive verb presents an overt subject or not. Thus, using the traditional case labels, we note the following examples from Comrie (1978) (see Timberlake 1975 for a thorough discussion):

(44) a. *Maija-∅   söi kala-n*
       Maija-nom ate fish-acc
       'Maija ate the fish'
    b. *Syö   kala-∅ !*
       eat-imp fish-nom
       'Eat the fish!'

Similarly objects of impersonally used (i.e. subjectless) transitive verbs also take nominative objects. On the other hand, if the subject of the (affirmative) transitive verb is present, the nonpartitive object is 'accusative' ('antiergative' in Comrie's terminology, but in any event different from the nominative). So case marking of one DNP then is contingent on the presence or absence of another.

A different type of dependency between DNPs is illustrated in Dalabon (Australia; Silverstein 1976, cited in Comrie 1978): Here case marking is contingent on the 'chain of being' relations between the transitive subject and the object. Thus if the subject is animate (as in (45) below) and the object inanimate the subject is not case marked. But if both are animate the subject takes an ergative marker:

(45) *Buluŋan-∅   ga'manbuniŋ*
     my-father   he-made-it
     'My father made it'

(46) *Buluŋan-   yi   wuduwud ga'nan*
     my-father- erg baby   he-looks+at-him
     'My father is looking at the baby'

A third related type of example is given by case marking in ergative languages when a 'normally' transitive verb is used 'absolutively' as *He wrote and wrote* as opposed to *He wrote the letter*. Some ergative languages cited in Comrie (1973) use an absolutive ending for the absolutive use, and others, like Hindi use the ergative. In the former case then, case marking of the subject is not predictable from the identity of the verb.

It appears then that the domain of case marking must not only include the DNP and the VP or TVP but also other DNPs in the clause. And yet we have only begun to touch upon the other properties of the sentence to which case marking must be made sensitive. Thus in many languages (Hindi, Georgian) case marking is sensitive to Aspect. Subjects of perfective sentences are marked in one way, subjects of imperfectives another. And since the same verb can occur in many different aspects this property of case marking is not predictable from the verb.

Further, it is quite common that case marking of DNPs is different according as the sentence they are in is negative or not. Thus, as is well known, DOs in Russian and Czech may be genitive if the sentence they occur in is negative. This is also true for Finnish and Lithuanian, as Moravcsik (1978) points out, and further the same holds for certain intransitive subjects. Compare (47a,b) from Moravcsik:

(47) a. *proexal avtomobil'*　　(Russian)
　　　 went-by car (nom)
　　　 'A car went by'
　　 b. *ne proexalo aytomobilja*
　　　 neg went-by car (gen)
　　　 'Not a car went by'

Analogous claims again hold for Finnish and Lithuanian. Worse, in Kawaiisu (S. Numic, Uto-Aztecan; see Munro 1976 for details) Subjects of negative sentences are marked (overtly) like DOs, whereas in nonnegative sentences they are not overtly case marked. In fact, Subjects of embedded clauses generally take object-marking. So it may be the case that case marking is not independent of whether the sentence the DNP occurs in is a main clause or an embedded clause. In Wappo (Li and Thompson 1975, cited in Gary and Keenan 1977) the overt case marking on Subjects is simply lost in finite subordinate clauses:

(48) a. *ce kew-i ew toh- ta?*
　　　 that man-Subj fish catch-past
　　　 'That man caught a fish'
　　 b. *Ah ce kew-Ø ew toh- ta? hatiskhi?*
　　　 I that man fish catch-past know
　　　 'I know that that man caught a fish'

It appears then that case marking is fully clausal in nature. And on reflection this is not surprising. In very many languages pre- or postpositions not only mark DNPs, they also mark various types of subordinate clauses. Thus subordinate clauses like *on arriving at the station, John went straight to the ticket counter; The meeting began upon John's arrival,* etc., may well be translated in heavily case marking languages by ordinary

morphologically bound case markers. Recall in this connection the wide-spread use of the 'absolutive' construction in the IE languages. (49) from Ancient Greek below is illustrative (diacritics omitted):

(49) *elthont-a   eis ten polin, ho didaskal-os edeixe ton paid-a*
     coming-acc   into the city, the teacher- nom showed the child-acc
     to sokrat-e
     the Socrates-dat
     'Coming into the city, the teacher showed the child to Socrates'

The nonfinite verb on the (subjectless) subordinate clause is case marked accusative, which indicates that the understood Subject of *coming* is coreferential with the DNP marked accusative in the main clause, namely the child. Had *coming* been marked nominative, it would have been the nominative DNP, *the teacher*, who was coming; and had it been marked dative it would have been *Socrates*. And had it been in the 'absolutive' form, the genitive case (ablative in Latin, Dative in Gothic, Locative in Sanscrit), the understood Subject of *coming* would have had to be different from any of the major participants in the main clause and thus already understood from the prior discourse.

The general point here is that case marking commonly has a variety of clausal and cross clausal functions. The domain of case marking then, even restricted to DNP marking, is properly the entire sentence in which the DNP occurs.

Note finally, that if case marking is a sentence level property we might expect reflexes of it to show up elsewhere than on the DNPs. And this in fact is correct, at least where case marking is understood broadly enough to cover markings on indirect objects and obliques. Recall the Latin paradigms in which 'prepositions' may show up either on the DNP, or on the verb, or both. Thus from a verb like *ferre* 'to carry' we may form derived verbs such as *inferre* 'to carry into', *transferre* 'to carry across', *ex+ferre* 'to carry from', *ad+ferre* 'to carry to', etc. The DNP 'Object' of such verbs may or may not also carry the preposition. Thus Caesar may carry the war into Italy, he may 'incarry' the war into Italy, or he may "incarry' the war Italy. Similar double markings occur in Hungarian and to a lesser extent Modern Russian. (50) below from Hungarian is illustrative.

(50) *Janos ra-Ø- te- tt- e.   a kalap-ot az asztal-ra*
     John on-it-put-past- 3sg3sg the hat- acc the table- on
     'John put the hat on the table'

Further, it is in fact quite common to find that that element which marks the relation a DNP bears to the action expressed by the verb may be marked exclusively on the verb. Such cases arise commonly in Caucasian languages, Penutian languages (Chinook, Totonac), Ancient Greek (see Keenan 1976 for examples), Arawakan languages (e.g. Machiguenga; see Gary and Keenan 1977 for examples) and Bantu languages. (51) below from Kinyarwanda (see Kimenyi 1976 for a thorough discussion) is illustrative:

(51) *Yohani y- a- andik-iish-ije-ho ameza ikaramu*
     John   he-past-write-inst-asp-loc table pen
     'John wrote on the table with the pen'

We may conclude then that syntactically speaking case marking is a function taking the entire sentence as its domain and thus the only prediction made by the FDP is the correct one, namely that the form of case markers may vary with the properties of the sentence it applies to.

3. *The Left-Right Order Relation.* The function-argument structures concerning CNPs and DNPs distinguished in *Logical Types* turn out to support, quite unexpectedly, another correlation between SF and LF: Namely, the characteristic left-right order of the constituents which express these structures. Thus:

THE SERIALIZATION PRINCIPLE (SP)

Different functional expressions taking the same class of argument expressions tend to serialize on the same side of their argument expressions.

THE DISSIMILATION PRINCIPLE (DP)[2]

Functional expressions taking DNPs as arguments and functional expressions taking CNPs as arguments tend to serialize on the opposite side of their argument expressions.

The SP predicts that VPs, TVPs, PossPhs, and Adpositions (= Prepositions in English) will all occur on the same side of their DNP arguments. Similarly, APs, Relative Clauses, Articles, Determiners, and Quantifiers will tend to occur on the same side of their CNP arguments. And the DP says that these two sides are different. Thus there should only be two characteristic word order types with respect to functional expressions of DNPs and CNPs: Either Function + DNP and CNP + Function or DNP + Function and Function + CNP. Although these predictions are too strong in one respect and so will be weakened slightly below they do make a very substantial number of correct predictions as we shall show and thus do constitute evidence in support of a general correlation between SF and LF.

Thus consider the characteristic word order patterns in Verb Final (SOV) languages, easily the most widespread word order type across different geographical areas and genetic groupings. The relevant word order correlations are:

(52) a. Subject + VP (e.g. John sings; John Mary kissed)
   Object + TVP
   DNP + Postposition (e.g. the garden-in)
   DNP + PossPh (e.g. John's father)
  b. Adjective + CNP (e.g. tall man)
   Relative Clause + CNP (e.g. the apple eating man)
   Article + CNP (the man; this man)
   Quantifier + CNP (every man)
   Numeral + CNP (two man)

Clearly all the functions on DNPs occur on the same side of the DNP, the right, and all the functions on CNPs occur to the same side, the left. So the SP is fully supported. And these two sides are different, so the DP is

fully supported, at least as far as our current knowledge of the word order correlates goes. Specifically, we know more about some of the word order patterns than we do about others. In particular, as indicated earlier, the internal construction of numeral expressions is complex, nonuniform, and not well known across languages. Further, in all cases here, and below, it should be recognized that we need large scale studies of the word order correlates to make our claims more accurate, and the few such studies we have (see e.g. Hawkins 1979a and 1979b) will surely show that the internal analysis of word order patterning is more complex than indicated above. Nonetheless, relative to the current state of our ignorance, the above mentioned correlates are (with the serious hedge regarding numerals) the best gross statement we have for the structures referred to. (Other correlates, such as the position of Auxiliaries and Main Verbs are not considered here.)

With these qualifications then, consider the second word order type predicted by the SP and DP. It would be:

(53) a. VP + Subject
         TVP + Object
         Preposition + DNP
         PossPh + DNP (e.g. the father of John)
      b. CNP + Adjective
         CNP + Relative Clause
         CNP + Determiner (= Article, Quantifier, or Numeral)

The first two claims in (53a) predict that a language of this type will have VOS as a basic word order. In fact, such languages are rare, although about 12 such cases are known. See Keenan (1978) for a study of their typological properties. On the other hand, though the VOS type is rare, the predictions re word order above are correct (modulo the hedges given above for SOV languages).

But what about the word order types, SVO and VSO not predicted by the two principles? Not only do they exist, but they are both more widespread than the VOS type, though very significantly less so than the SOV type. (54) below summarizes the known word order types characterized in terms of the realtive order (where fixed) of Subject, Object, and Verb:

(54) SOV > SVO > VSO > VOS > OVS

The chart below gives a rough breakdown of the distribution of word order types across the major genetic-areal grouping of languages. A major heading is understood to mark a single genetic group unless indicated otherwise. Subheadings mark genetically coherent subgroups.

GENETIC-AREAL DISTRIBUTION OF WORD ORDER TYPES

1. *Australian*
    SOV with much freedom of word order

2. *Indo-Pacific* (6 genetically independent groups)
    SOV

3. *Austro-Tai*
   SVO and VSO both prominant orders. (Some SOV in New Guinea)

4. *Austro-Asiatic*
   4.1 *Munda*—SOV
   4.2 *Mon-Khmer*—SVO

5. *Sino-Tibetan*
   SOV

6. *Dravidian*
   SOV

7. *'Boreo-Oriental'* (possibly four genetically independent subgroups)
   7.1 *Altaic* (Turkic, etc.)—SOV
   7.2 *Uralic* (Samoyedic and Finno-Ugric)—dominantly SOV with some SVO
   7.3 *Japanese, Okinawan, Korean*—SOV
   7.4 *Paleo-Siberian*—dominantly SOV

8. *Indo-European* (extant groups)
   8.1 *Baltic*—SVO
   8.2 *Slavic*—SVO
   8.3 *Germanic*—SVO with traces of SOV
   8.4 *Italic*—SVO with hints of VSO
   8.5 *Indo-Iranian*
      8.5.1 *Iranian*—SOV
      8.5.2 *Indic*—SOV
   8.6 *Dardic*—SOV
   8.7 *Greek*—SVO
   8.8 *Albanian*—SVO
   8.9 *Armenian*—(modern) SOV
   8.10 *Celtic*—VSO

9. *Caucasian* (perhaps 2 genetic groups)
   SOV dominant, scattered SVO

10. *African* (4 genetically independent subgroups)
    10.1 *Khoisan*—SOV
    10.2 *Nilo-Saharan*—VSO, SVO, and SOV all attested
    10.3 *Niger-Kordofanian*—SVO
    10.4 *Afro-Asiatic*
       10.4.1 *Semitic* (modern)—SVO
       10.4.2 *Berber*—VSO
       10.4.3 *Chadic*—SVO
       10.4.4 *Cushitic*—SOV

11. *Amerindian* (13 genetically independent subgroups)
    SOV probably the dominant order (e.g. Eskimo-Aleut, Na-Dene (Athapas-
        canl, Hokan, Aztec-Tanoan, Andean-Equatorial, Macro-Chibchan)
    VSO heavily attested: (Salish, Wakashan, Penutian, Oto-Manguean)
    SVO is only rarely attested
    OVS is known in one case (Hixkaryana, a Carib language in Brazil,
        see Derbyshire (1977) for discussion)

It is clear from the table that SOV is clearly the dominant order among the world's languages. The relative proportions of SVO vs VSO are hard to evaluate given that I cannot accurately assess just how widespread it is among the Amerindian, though it is clear that it is a dominant order in the genetic groups indicated above. So SVO occurs as a dominant order in several Amerindian phyla, it is one of the dominant orders in Austro-Tai (Polynesian, Philippines, Formosan), and occurs as the dominant order in genetic subgroups of Indo-European (Celtic), Nilo-Saharan (Nilotic), and Afro-Asiatic (Berber). As well it was likely the historical order of older Semitic languages: Classical Arabic, Biblical Hebrew, Babylonian Aramaic. SVO on the other hand, is the dominant order in Europe (ignoring Celtic) though in the phylum as a whole, SOV is a very significant order (Armenian, Dardic, Iranian, Indic). In addition SOV is the dominant order in the extinct IE languages (Hittite) and is the most likely reconstructable order for Proto-IE. SVO is the dominant order in Niger-Kordofanian, a very large group, and occurs as a dominant order in many Austro-Tai language groups (Tai, Indonesian, Javanese). It is the dominant order in Chadic and modern Semitic and occurs sporadically in groups that are otherwise dominantly SOV. I consider it more widespread then than VSO.

VOS only occurs rarely, and then mainly in Phyla which independently present Verb Initial as a major order--Austro-Tai and Amerindian. OVS as noted is cited for only one case, though others related to it may have the same order. We shall not further consider this case here.

What about the word order patterns in SVO and VSO languages other than those involving the subject? Essentially they pattern like we predict for the VOS languages, though more 'doubling' (Hawkins 1979a) occurs, especially in SVO languages (e.g., such languages may have both prenominal and postnominal possessors, as English; both prenominal and postnominal adjectives, as French, etc.). Further and large scale study would be needed to justify that articles and demonstratives occur significantly more frequently postnominally than prenominally, as they are generally considered to do (e.g. Indonesian, Yoruba, Batak, etc.). These observations suggest the following condition:

SUBJECTS FRONT (SF)

skip two lines The Subject occurs to the left of other major constituents of Ss. Thus SF is maximally satisfied in both SOV and SVO languages, less satisfied in VSO languages, and still less in VOS and OVS languages. Note further that the SF is a much more specific principle than either the SP or the DP, as the latter quantify over classes of functional expressions, whereas the SF only mentions one specific argument expression--the Subject. If we take then the possible word order types for human language to be those obtained from the set defined by the general principles SP and DP, namely SOV and VOS, and modified in the direction of satisfying Subject Front we obtain (pending of course a more rigorous formulation) a surprising number of correct preditions:

(i) The possible types are SOV, VOS, VSO, and SVO. Just the well attested cases, with the exception of the one OVS language cited.

Note what a strong prediction (i) in fact is. There are 9 function-argument structures considered and thus in principle $2^9 = 512$ possible word order types. Of which only four are predicted to occur, and they do.

(ii) The principles predict that the SVO, VSO, and VOS languages will evidence the same serialization in their function-argument structures with the exception of the placement of the Subject. But all the others should be the same, and this is correct.

(iii) The combined effect of the three principles would appear to correctly predict the observed genetic-areal distribution of the word order types.

Thus the rigorous SOV type satisfies all the principles and is in fact the most widely distributed word order type. SVO fully satisfies Subject Front and fails the Serialization Principle (and the Dissimilation Principle) at just one point: the Subject should follow the VP, that is the TVP + Object. VSO also fails the SP (and thus the DP) at this point, and in addition satisfies Subject Front to a lesser extent than the SVO type and so is less widely distributed. Independently VSO yields a less good correspondence with LF since the TVP + Object does not form a constituent in SF. Finally VOS (and OVS) maximally fail Subject Front and might then be expected to have the least wide distribution.

Despite the logical looseness of prediction (iii) above it does seem to me that the approach we have taken to the word order correlates more successfully relates them to function-argument structure than earlier attempts. Vennemann (1972) is perhaps the first (but see note 2) attempt to systematically correlate word order types with function-argument structure. His approach however is not based on an independently justified logic. He takes for example DNPs as functions with VPs and Prepositions as arguments, and their outputs are supposed to somehow have the same category as the argument. But no natural logic treats, for example, *in* and *in the garden* as having the same type of denotation. Further, Vennemann's approach must rule out the Subject-VP relation in principle, so predictions like (iii) above, however loose, cannot in principle be stated in his framework.

More recently Gil (1979) relates function-argument structures to word order types in the framework of a prosodic theory. His approach assumes our SP and has an analogue of the DP, though much more elaborate. The two approaches are sufficiently different to resist quick comparison, but it is noteworthy that his approach does yield several predictions concerning distinctive prosodic structures in SOV and SVO languages, one of which (greater consonant/vowel ratios in the SOV group) he rigorously supports.

To conclude this section, we might wonder whether the Subject Front principle is anything other than an ad hoc device to get the word order types correctly predicted. Is there any independent reason to expect this principle, and is there any independent justification for it? The answer to both questions I feel is yes, although more work than can be presented here would be needed to justify this claim. So briefly:

As regards motivation, the Subject-VP constituent is the major function-argument structure of simple sentences and might well be expected to have

more perceptual or cognitive salience than the function-argument structures embedded within either the Subject or the VP. Thus, recall from the Meaning-Form Dependency Principle that how we interpret VPs may vary with the choice of Subject. Thus if the VP preceeds the Subject, hearers will have to suspend an exact interpretation for it until the Subject is enunciated. For example, if English were a VP first language there would be sentences beginning with *is flat . . . , is strong . . .* , etc., where the exact sense of the VP could not be interpreted until we knew whether the speaker was talking about roads, beers, voices, or whatever. Further, the problem would quite obviously be more serious in a VOS language than in a VSO language, since more material in a VOS language would be present before the Subject was enunciated. To say for example in a VOS language that *nobody loves both his father and his mother,* we would begin with, *loves the father of his and the mother of his . . .* , without being able to identify the referent of *his*. And if the language allowed sentential objects (the most likely position for complements of verbs of saying) we could begin sentences like <u>thinks that hurt himself he</u> . . . before the Subject phrase was enunciated, requiring the hearer to 'hold in temporary memory' *himself* and *he* until their referents can be established by the Subject. (See Keenan 1978 for further discussion of the disadvantages of placing the Subject at the end of the sentence.)

The motivation suggested above for presenting Subjects to the left of the other major constituents in a sentence does in fact suggest some independent motivation for Subject Front. In the first place, in those languages which do use a VOS order, we in fact never find that the Object position can be filled in surface by a full sentence. To say in those languages (e.g. Malagasy, Fijian, Toba Batak, Ineseño Chumash, etc.) things like *John thinks that Fred will win* a variety of alternatives are available, one of which must be used, and all of which present the thinker before the embedded clause. Thus the language (e.g. Malagasy) may sinply use a VSO order here, in violation of the normal word order. Or it may passivize (Malagasy, Toba Batak) yielding an order like *thought by John that will win Fred*. Or it may utilize any of various devices (clefting, topicalization) for presenting the Subject before the VP.

In the second place, all VSO languages and all VOS languages present means of fronting the Subject. Some of these may be clearly marked structures, e.g., a particle may be inserted after the fronted Subject and before the remaining VP, but such options always exist and, in all cases where I have relevant data (Malagasy, Fijian, K'ekchi (Mayan) and Toba Batak) such fronting devices are more commonly used than, for example, the 'backing' devices (e.g. Right Dislocation) in SOV or SVO languages.

In the third place, even in VSO languages, there will be a variety of ways to form complex (intransitive) verb phrases, e.g. by conjoining two independently intransitive verb phrases. The expected order then would be *sing and dance John*. And sometimes this order can be realized (Malagasy, Isthmus Zapotec, Tamazight (Berber)), but in general it seems more common to break open the VP yielding 'dissonant' structures like *sing John and dance*, where we have a surface coordination between a VP and a full Sentence. And if the VP is a TVP plus Object it is even harder to maintain the VP-Subject order, despite the highly dissonant structures that result.

That is, we would expect to get [*kissed Mary and danced*]$_{VP}$ *John*. (And sometimes we do, as in Malagasy.) But the more usual order is [*kissed John Mary*]$_S$ *and danced*.

These cases do then provide independent evidence for <u>Subject Front</u> as a constraint in Universal Grammar on the class of possible word order types.

4. *Theoretical Significance of Correlations between SF and LF*. Why, as linguists, should we be interested in establishing a correlation between the properties of SFs and those of LFs? The answer lies in our goal of trying to represent the linguistic competence of speakers of a language. To know, and thus to learn, a language is to know (learn) much more than simply what the class of grammatical SFs are. Even children from the youngest age do not <u>randomly</u> produce approximations to well formed SFs. Rather they use them with logical (and other) effect. That is, they use them to assert--to make true statements (or statements they intend as true). They use them to deny, and they readily make inferences from what others have said. In other words, in practice, speakers use SFs as LFs. It makes sense to say that what someone has said entails something else, or that it contradicts what someone else has said.

If we represent at least the logical properties of an SF by a set of LFs we may say that what someone knows when he knows a language is a set of pairs $<s,t>$ where s is an SF and t is an LF which represents one of the meanings of s. We may represent the entire language as the set of pairs $<s,t>$ such that s is in SF, t is in LF and t represents one of the meanings of s. (So the language will be a very proper subset of SF x LF.)

(To be sure this idealization abstracts away from much else we know when we know a language, such as most everything we might call 'sociolinguistic', but for the moment this abstraction is sufficient for our purposes.)

To know a language on this view then requires that we know what the class of SFs is, what the class of LFs is, and most important that we know which SFs are paired with which LFs. This last claim can be represented by saying that the ideal speaker knows (in the sense of having internalized) a function which associates each SF s with the set T_s of LFs, each of which expresses one of the meanings of s.[3] In fact, one of the explicit goals of current linguistic theory (at least the Revised Extended Standard Theory, EST) is to define such a function, although work specifically on that topic has in fact been rather scant.

Thus, to account for the linguistic competence of the (ideal) native speaker we shall want to characterize the set of SFs, the set of LFs, and the interpreting function which relates the two. And we shall want to use this characterization to explain what it is that people learn when they learn a language and how it is that children, with a limited and imperfect exposure to the language, nonetheless learn to identify a substantial subset of the pairs (s,t) which constitute the language.

Part of the explanation for this latter fact within generative grammar

has been that humans come biologically equipped with a certain syntactic potential. So the child, or anyone else, is not prepared to accept as a possible set of SFs a very large number of the sets of expressions which can be formally defined. We rightly do not expect that somewhere in the unexplored wilds of New Guinea we will run across a language all of whose sentences have prime length. Thus much work in generative grammar has, rightly to my mind, been concerned to constrain the class of possible grammars the learner can potentially accept. These general constraints would then directly reflect our innate syntactic potential.

Analogous claims hold for the class of possible interpreting functions. We cannot believe that the way humans associate meaning with form, or LFs with SFs, is random. Given a set of SFs for a language, not just any function into a (power) set of LFs is a possible interpreting function for that language. Imagine for example a function F which, for every surface form s, maps s onto a set of LFs none of which mentions any of the LFs which F associates with any of the constituent parts of s. So the LFs which F might associate with *John came early and left late* could be just those that a 'right' interpreting function associated with *Not all unicorns have horns*. Mathematically it is easy to construct such nonadmissible interpreting functions.[4]

So to characterize our competence we want to define the class of possible interpreting functions, and to explain (at least in part) the learnability of human languages, we want to define the class of possible human languages, that is the possible sets of pairs (s,t), as narrowly as possible. Clearly the most constrained class of interpreting functions is the set of identity functions. That is, for any language L, the set of LFs for L would be identical to the set of SFs for L. This would directly explain why speakers can use SFs with logical force (assert, deny, infer, etc.).

Unfortunately this identification does not seem possible on any theory of SF ever propounded. The reason is that in all theories, at least some SFs will be semantically ambiguous and thus correspond to more than one LF. Note that the objects on which the entailment relation is defined (LFs) cannot be ambiguous in the sense of having more than one truth value in some state of affairs. If p were such a formula then we could argue both that p entails p and that p does not entail p, which is to say that the entailment relation itself is not well defined. (To see this, note that given the existence of a state of affairs in which p is true on one reading but false on another it follows that p does not entail p. On the other hand, 'for any state of affairs M, if p is true in M then p is true in M' is itself trivially true. Which is to say p entials p.)

But how badly must the identification of SF with LF fail? Which is to say, how ambiguous are SFs? The question has no obvious answer since it depends at least on what SFs our theory defines, and as we have seen, that is a complicated matter. Still, one kind of SF which is semantically ambiguous in all theories are those containing semantically ambiguous lexical items of the same grammatical category and subcategory. Thus *John is a bachelor* we are told is four ways ambiguous according as John is a male seal, a knight's helper, etc. And since it can be true on one reading and

false on another, we will need distinct LFs to correspond to each of these readings. Notice however that each of those LFs could be structurally isomorphic to the SF. They might just differ in that within the logic would be four distinct logical common nouns, *bachelor₁*, *bachelor₂*, etc. This type of ambiguity then, while serious, still permits a very strong form of correspondence between SF and LF. Namely, identity up to lexical disambiguation.

But of course linguists and philosophers have been much concerned with nonlexical ambiguities. Many of these however will be eliminated in a reasonably rich theory of SF. Thus most theories would assign *flying planes can be dangerous* two distinct surface forms. (On one, *flying* would be the head of the Subject, and on the other it would be a modifier. On one the head would be marked + plural, and on the other not, etc.) And it appears that many of the classical structural ambiguities are in fact not ambiguous at a level of SF. Note that even a subcategory difference in a lexical item yields a different labelled node in SF. Thus very possibly *the shooting of the hunters* can be naturally assigned to SFs according as the underlying verb *shoot* is transitive or a derived intransitive.

It is worth noting in this regard that elaborating the base, as by proliferating the number of categories and subcategories, does not seriously complicate the grammar as a whole. That is, the class of context free or even context sensitive languages is a very constrained class. If all of English could be given in a context free (sensitive) form the learnability problem would be solved. On the other hand, the class of languages generable by context free bases plus length decreasing filters is enormously larger. Thus to constrain the class of grammars we have an interest in enriching the base and limiting as much as possible the filters used.

On the other hand, there are several sorts of 'possible' ambiguity which have been studied by philosophers and linguists alike which are much less likely to be represented by distinct SFs. The major cases known to me are the following: (1) cross-reference ambiguity, as in *John told Bill that he was clever*; (2) Quantifier scope ambiguity, as (supposedly) in *Every man loves a woman*; and (3) transparency/opacity mabiguity as in *Suzie wants to marry a Swede* or *John believes that the author of Waverly was the author of Waverly*.

Just how far from isomorphism these cases will force the interpreting function is unclear at the moment. A serious part of the reason is that we have no really clear pretheoretical intuition of what shall count as semantic ambiguity.[5] Thus the position that the *he* in *John told Bill that he was clever* is simply freely generated and its interpretation depends on the context of utterance is at least defensible. So on that view the sentence would not be logically ambiguous. Its 'pragmatic' potential would have to be explicated within a theory of pragmatics or speech usage. However, while many are working in this area, there is not yet, to my knowledge, a sufficiently developed body of science here to be called a theory, and the claim that our sentence is only pragmatically one way rather than logically one way is only a possibility. Similarly one might want to sweep quantifier scope ambiguity under the rug of pragmatics, and reduce the transparency/opacity ambiguity to one of scope of variable

binding operators (an approach which has been reasonably successful for some of the classical opacity cases; but many unsolved problems remain), thus reducing it in effect to the quantifier scope cases.

So it appears . .thinkable that there could be, at least in principle, something like an isomorphism between an SF and any of its LFs. However, it seems more likely that the mismatch between SF and LF will be greater than that. The ambiguity issue only stated one in principle condition that required a mismatch. But just because a SF is semantically unambigu-ous it will not follow that, say, its constituent structure corresponds point for point to that of its LF. And further, on some approaches to SF, many surface forms would be generated which would correspond to the null set of LFs. Examples might be *the man was hitting themselves, John loves each other*, etc. So on this view the interpreting function itself would in effect act as a kind of filter on the SFs.

So it seems to me then that the best assessment of the relation between SF and LF lies somewhere strictly in between isomorphism and randomness. How close to one pole or the other is an open question, but an important one, since our answer to the problem will determine how successfully we can characterize our linguistic (semantic + syntactic) competence. We should, it seems to me, be looking for theories of SF and LF in which the correspon-dence is as close as possible, while still accomodating the ambiguity 'facts'. For the closer the correspondence the more we have accounted for a speaker's competence in making inferences, denying, etc. Further, a close correspondence theory gives us perhaps an additional means of explain-ing how children learn quickly with limited exposure. It not only limits the ways the child may attempt to assign meaning to form but it gives an additional reason for the child's quick learning. Namely, we may assume the child is motivated both to express his needs (and in fact his under-standing of the world) and also to understand what others say. That is, there is an intense motivation to be able to perform and comprehend 'semantic' acts. So if SFs are directly useful in coding and in decoding meaning the child is motivated (or at least pushed) into learning the semantically significant properties of SFs. In more concrete terms, the child is motivated to learn the grammatical distinctions between common nouns and intransitive verbs if these distinctions are useful in expressing meanings and understanding the meanings others communicate. If on the other hand, that distinction had no semantic value the child would be less motivated to learn it, and so presumably slower in acquiring it. The lazy child might relapse into conjugating nouns or putting determiners on verbs. We know in fact from psychology (though the cases studied are not directly comparable to the case we are considering) that people learn meaningful material more quickly than meaningless material. To the extent that SFs can be directly interpreted, that is, treated as LFs, the distinctive properties of SFs are meaningful and can be expected to be learned more quickly than if these properties were meaningless.

5. *Conclusion.* Unfortunately, it seems to me, much of the post-*Aspects* work in generative grammar has, in different ways in different cases, moved away from a conception of SF in which it is closely related to LF.

The original conception of syntactic theory in *Syntactic Structures* however asemantic in conception, lent itself very nicely to a close relationship between LF and syntactic form generally. The kernal sentences on that view represented a <u>finite</u> amount of information a language learner had to learn by brute force. The unbounded set of derived structures however could naturally have their meanings represented as a function of what they were derived from since the generating functions (the transformations) were in a naive sense semantically interpretable. Thus a child could learn the meaning of Yes-No Question Formation (one of the generating functions) by learning how the meaning of a small sample of questioned kernal sentences differed in meaning from the kernal sentences they were derived from. Then the meaning of any yes-no question can be ascertained relative to the declarative it is formed from simply by recognizing that it is, syntactically, a yes-no question. Analogous claims held for most of the major transformations in *Syntactic Structures*. A proper semantic theory at this time would have 'simply' faced the task of indicating how the output of a transformation varied in meaning as a function of its input. And this is the sort of thing that logical theories do. Logical derivations by conjunction, disjunction, quantification, and negation do not preserve meaning. Rather we state the meaning of the output as a function of that of the input. Needless to say of course the semantic operations of English would have been at a higher level of complexity than in Classical Logic, but not in principle different.

But as is well known linguistic theory did not develop in this way. By 1964 and 1965 assigning meaning representations to SFs was a recognized goal of the theory, but the theory was reorganized in such a way that only base structures and lexical items were to be semantically interpreted. Taking off from the point, Generative Semantics attempted to argue that the putout of the Base differed very significantly from SFs, and that in fact the base structures (modulo a few differences as regards quantified expressions) resembled the LFs of the classically given first order logic (CL)! This is a terribly negative view of the relation between SF and LF, although that was perhaps not so apparent at the time.

But consider that the entire syntax of LFs in CL can be given in a single paragraph. Here it is: CL has a denumerable set of variables, x_1, x_2, etc. each having the category of Name. It has a set of function symbols each of unique degree $n \geq 0$, and a rule which says that the concatenation of an n-place function symbol with n occurrences of names is a name. The language also has a set of relation symbols each of unique degree $n \geq 1$, and a rule which says that the concatenation of an n-place relation symbol with n names is a formula. And it has two rules for deriving complex formulas. One says that if S and T are formulas then (not S), (S and T), and (S or T) are formulas. And the other says that if S is a formula and x a variable then (for every x, S) and (for some x, S) are formulas. And that's all.

Now the best generative grammars of English we have (Stockwell, Schachter, Partee 1973) run into hundreds of pages, and they are clearly inadequate. Assuming that the meanings expressible in English can be done in CL or a slight extension thereof, we are saying that almost all of English syntax is needless as regards the expression of (logically

representable) meaning. That is, we could have done with a one-paragraph grammar rather than the massive thing we apparently require. And since the diversity of structure present in SFs vastly exceeds the primitive syntax of CL, this view forces a massive correspondence failure between SF and LF. I find this a discouraging view, a kind of modern Babel theory in which Man is punished by God by having to use an absurdly complicated grammar when he could have done with a sleek one-paragraph one.

Nor does this view conform to my experience as one who has been concerned to represent the logical properties of natural language. The more I have looked at natural language the more I have found that almost everything in surface is relevant to the logical interpretation of SFs.

One might expect on the other hand that there would be a more optimistic view of the relation between SF and LF within EST syntax. It is clearly a goal of that theory to map SFs onto LFs, and the closer the correspondence between the two the easier it will be to define the mapping. But I am uneasy about the relation between SF and LF within this theory. On the one hand, in 'Questions of Form and Interpretation' Chomsky (1974: 16) espouses the view that there is a close relation: '. . . the thesis of "absolute autonomy formal grammar" . . . would not imply that there are no systematic connections between form and meaning. No one, I am sure, has ever doubted that there are highly systematic connections and that a major problem of linguistic theory is to determine them.' On the other hand, Chomsky and Lasnik (1977:429) assert that they believe, 'There is . . . some empirical support for the belief that the syntax of LF is close to that of standard forms of predicate calcalus . . .' (though that assumption does not appear to me to play an important role in the article).

Given our earlier discussion that an assignment of LFs to natural language expressions is nonobvious in the same sense in which assignment of SFs is, it is certainly unwarranted to assume that LFs have one or another particular shape. To my knowledge no specific attempt has been made within EST syntax to define a function mapping SFs onto descriptively adequate sets of LFs drawn from Classical Logic. We cannot then assess the feasability of defining such a function nor the demands it would make on the language learner.

I can think of only two reasons why, as generative grammarians, we might want to assume prior to an investigation that the LFs of CL are an appropriate range for an interpreting function for natural language SFs. I shall argue here that these reasons are in fact only apparent.

The first is that according to a certain well defined metric of logical complexity (the degree of recursive unsolvability of the predicate logically true) CL is reasonably simple, though not absurdly so. Thus in CL logically true is a recursively enumerable predicate, whereas in second or higher order logics this is not the case. Thus for example in second order logic there can in principle be no complete syntactic characterization of the entailment relation. So if the child (or whoever) only has to know things of first order difficulty, he has to know less than if he has to know things of second or higher order difficulty. This point is surely correct. To claim that the natural logic for English has the full

expressive power of second order logic is to make a stronger claim about our logical competence than to claim that first order logic is sufficient.

However, the order of a logic is not tied to any particular syntactic instantiation, such as that in CL. The order of a logic is strictly determined by the set theoretical type (relative to the universe of discourse) of the variables in the logic. What makes first order logic first order is that we can only quantify over members of the universe of discourse U. In second order logic we can quantify over subsets of U, and so we can say things like 'For every object x there is a property P such that P(x)' and so on. So a logic which only allows quantification over U is first order, and that commits us to very little concerning the class of LFs the logic defines. It only requires that we have variables of a certain type but otherwise we are free to construct the rules defining the LFs as we please. For example, the extensional logic proposed in *Logical Types* is in all essential respects a first order logic. (The range of the variables in that logic appears to be sets of higher types, but in fact they are so constrained as to be in a one-one correspondence with the members of U.)

Secondly, the question of close correspondence between SF and LF is in principle independent of the question of autonomy of syntax. I take the autonomy thesis to be essentially that our syntactic capabilities are innately determined up to some level of specificity and that these capabilities are independent of other things, such as the semantic ends we make them serve. In fact construed in this way the autonomy thesis seems to me clearly true. There is to be sure a serious empirical matter as to how much of say English syntax is innately determined and how much represents a 'free' choice within the bounds of our innately determined possibilities. But the in principle question seems clear: only humans use language, so what more natural explanation than that humans differ biologically in this respect from other mammals?

But from the autonomy thesis alone nothing <u>follows</u> concerning the relation between SF and LF. It certainly does not follow for example that agreement pairs do not correspond to function-argument structures in the way we have claimed in this paper. Nor will it follow that certain function-argument structures are not commonly expressed by SFs which exhibit regular left-right order relations among their parts. Nor will it follow that the interpretation of complex SFs is completely unrelated to the interpretation of their constituent parts.

Whether correspondences exist between SFs and their LFs is an empirical matter. To say that there is a massive correspondence failure is to say that man has made very poor use of his syntactic endowment. It is to say that man makes very many syntactic distinctions which are in fact irrelevant to the expression of meaning; and that he fails to code in his syntax distinctions which are semantically relevant. I can believe that we have not made perfectly efficient use of the syntactic instrument we come endowed with; that idioms for example do present meaningful parts whose meaning is (synchronically) irrelevant to the meaning of the entire idiom. And I can believe that there may be certain systematic ambiguities in language, such as the opacity/transparency ones cited earlier, which are not regularly coded in SF.

But I cannot believe that the syntactic elaboration present in each language can be collapsed into the pitiful syntax of CL. How will we represent the voicing systems of the Philippines? The tense systems of Bantu? The switch reference systems of New Guinea and the Americas? Where shall we put the deixis systems present in all languages? Where will imperatives and hortatives go? And subjunctives? And gerunds? What about the rules which most if not all languages have which convert Ss and VPs into NPs and Adjectives? And Adjectives into abstract nouns? Where to put the mass nouns, and noun compounds? And do we really not logically distinguish common nouns from intransitive verbs? And adjectives from adverbs?

Surely if we have all this syntax in our heads we will use it for some semantic effect. It may to be sure be difficult to explicate just what the semantic effect is in particular cases, but rather than accept a punishment theory--all that syntax is useless--I would recall the German proverb often cited by Einstein: 'The Lord is subtle but he is not mean'.

NOTES

[1]Several alternatives for the analysis of PPs are in fact suggested in *Logical Types*. The version fully presented there was one in which Prepositions combined with either intransitive verbs (VPs) or transitive verbs (TVPs) to extend the class of DNP arguments they take. The analysis in which this semantics is presented with Prepositions as functions taking DNPs as arguments has been worked out by Faltz (Dept. of Linguistics, UCLA). For simple DNPs that analysis is equivalent to the one in *Logical Types*. If the DNPs are properly quantified there may be relative scope differences between the DNP object of the Preposition and the Direct Object of transitive verbs.

[2]It is interesting to note that perhaps the first scholar to investigate word order correlates, Abel Bergaigne, in 1875, arrives at a very similar analysis: they can be expressed by a 'double formule: le terme qualifiant précède le terme qualifié et le terme régi précède le terme régissant'. The quote is cited in Holland (1976:413).

[3]We need not require that T_s be the entire set of LFs each of which expresses one of the meanings of s. It is sufficient that T_s be a set of LFs such that any logical form which represents a meaning of s is logically equivalent to one in T_s. T_s then can be expected to be finite for any s.

Note incidentally that whether the set of pairs $<s, T_s>$ is recursive is completely independent of whether an arbitrary chosen logical form is valid, that is, true in all interpretations of the logic.

[4]For example, let I be a correct interpreting function for a set of SFs. So for each SF s, I(s) is the set of descriptively adequate LFs for s. Enumerate the set of SFs (so for each SF we may associate a unique natural number n). Pick any wierd numerical function, say the function f which maps each natural number n onto $100 + n^3$. Define a 'wrong'

interpreting function I* which maps each surface form s onto $I(\mathcal{L}(n))$,
where n is the number in the enumeration of s, *and for the proper form whose
number is $f(n)$*

[5]For example, as regards entailment we do have a good pretheoretical
intuition. Namely, an (unambiguous) sentence S entails an unambiguous
sentence T iff T is true in all the cases in which S is true. Model
theoretic semantics is an attempt to formalize this notion by formally
representing the notion 'true in a state of affairs'.

The best pretheoretical intuition behind the notion 'ambiguity' which
I can come up with is at best only a necessary condition for a sentence to
be ambiguous. Namely, if S is ambiguous between meanings A and B (how-
ever represented) then a speaker on a literal, sincere, etc., use clearly
intends either A or B. Thus the speaker cannot respond 'I don't know' to
the query 'Do you mean A, or B?' We may attempt to formally represent
this intuition by designing an interpreting function for SFs which maps
the pretheoretically judged ambiguous ones onto a set containing non-
logically equivalent LFs. But the descriptive adequacy of the mapping is
only as good as is our pretheoretical intuition of ambiguity, and that
in fact is not very good.

REFERENCES

Anderson, S. R. 1971. On the role of deep structure in semantic inter-
    pretation. Foundations of Language 7. 387-96.
————. 1977. On mechanisms by which languages become ergative. In Li
    1977, 317-63.
Capell, A. 1971. Arosi grammar. Pacific Linguistics Series B—No. 20.
    Canberra: The Australian National University.
Chomsky, N. 1974. Questions of form and interpretation. Manuscript.
    Department of Linguistics, MIT.
———— and H. Lasnik. 1977. Filters and control. Linguistic Inquiry 8.
    425-505.
Comrie, B. 1973. The ergative: Variations on a theme. Lingus 32.
    239-53.
————. 1978. Ergativity. In Lehmann 1978.
———— and S. Thompson. 1978. Lexical nominalizations. Manuscript.
    Department of Linguistics, UCLA.
Derbyshire, D. C. 1977. Word order universals and the existence of OVS
    languages. Linguistic Inquiry 8. 590-99.
Dixon, R. M. W. 1972. The Dyirbal language of North Queensland. Cam-
    bridge: Cambridge University Press.
Gary, J. and E. L. Keenan. 1977. On collapsing grammatical relations in
    universal grammar. Syntax and semantics: Grammatical relations,
    Vol. 8, ed. by P. Cole and J. Sadock, 63-121. New York: Academic
    Press.
Gil, D. 1979. Prosodic structure and linguistic universals. Manuscript.
    Department of Linguistics, Tel Aviv University.
Hawkins, J. 1979a. Implicational universals as predicters of word order
    change. Language 55.
————. 1979b. Two types of word order universals: Implicational and

distributiónal. Manuscript. Department of Linguistics, University of Southern California.

Holland, G. 1976. The shift from postposition to preposition: Evidence from Early Greek. Proceedings of the Second Annual Meeting of the Berkeley Linguistics Society.

Keenan, E. L. 1976. The logical diversity of natural languages. Annals of the New York Academy of Sciences 280. 73-92.

————. 1978. The syntax of subject-final languages. In Lehmann 1978.

—————— and L. Faltz. 1978. Logical types for natural language. Occasional Papers in Linguistics. No. 3. Department of Linguistics, UCLA.

Kimenyi, A. 1976. A relational grammar of Kinyarwanda. Ph.D. dissertation. Department of Linguistics, UCLA.

Lehmann, W. (ed.). 1978. Syntactic typology. Austin: University of Texas Press.

Li, C. (ed.). 1975. Mechanisms of syntactic change. Austin: University of Texas Press.

—————— and S. Thompson. 1975. Subject and word order in Wappo. Manuscript. Department of Linguistics, UCLA.

Montague, R. 1973. The proper treatment of quantification in ordinary English. Approaches to natural language: Proceedings of the 1970 Stanford Workshop on Grammar and Semantics, ed. by J. Hintikka, J. Moravcsik, and P. Suppes, 221-242. Dordrecht: Reidel.

Moravcsik, E. 1978. Ergative and accusative patterns. Manuscript. Department of Linguistics, University of Wisconsin.

Munro, P. 1976. On the form of negative sentences in Kawaiisu. Proceedings of the Second Annual Meeting of the Berkeley Linguistics Society, 308-19.

Murane, E. 1974. Daga grammar. Norman, Okla.: Summer Institute of Linguistics.

Partee, B. 1976. Montague grammar. New York: Academic Press.

Payne, J. 1978. Phrasal conjunction. Manuscript. The Typology Project, Department of Linguistics, UCLA.

Reinhart, T. 1976. On understanding poetic metaphor. Poetics 5. 383-402.

Silverstein, M. 1976. Hierarchy of features and ergativity. Grammatical categories in Australian languages, ed. by R. M. W. Dixon, 112-71. Canberra: Australian Institute of Aboriginal Studies.

Stockwell, R., P. Schachter, and B. Partee. 1973. The major syntactic structures of English. New York: Holt, Rinehart, and Winston.

Timberlake, A. 1975. The nominative object in Finnish. Lingua 35. 201-30.

————. 1977. Reanalysis and actualization in syntactic change. In Li 1977, 141-77.

Vennemann, T. 1972. Analogy in generative grammar: The origin of word order. Paper read at the XIth International Congress of Linguists. Bologna, Italy.