Semantic Categories and Syntactic Distribution

Yasuaki Abe

1. The Problems

The syntactic distribution of a particular lexical item can be determined to some extent by the syntactic category it belongs to. Given the syntactic category of an item, we can predict where in a sentence this item occurs. For instance, if we have a noun like *boy* in mind, we'll immediately know the local environments in which it occurs given the phrase structure rule of the following form.

(1) X → ...... N ......

Moreover, if there is a higher order phrase structure rule that defines the occurrence of X in Y, it becomes inevitable that N occurs in all the X positions in Y. Thus, if we take X in (1) to be NP and if we have the following phrase structure rules, an N like *boy* must occur in all NP positions in S.

(2) S → NP VP′
VP′ → Aux VP
VP → V (NP) (PP)
PP → P NP

(3) The boy will hit a boy for the other boy.

Thus, it is hard to imagine a certain lexical item that freely forms an NP occurring only in some NP positions of the sentence.[1] It would be very strange if a certain type of NP occupies, say, only the subject position and nowhere else.[2]

However, this sort of distributional asymmetry can actually be found in natural languages. The existence of such a phenomenon has often been discussed. Cf. Klima(1964), Lasnik(1972), and Postal(1974). Consider the following sentences, taken from Postal(1974).

(4) a. Not many gorillas have learned to tap-dance.
b. Not much sense can be made out of that proposal.
c. Not many farmers are easy to convince.

(5) a. *Joe kissed not many models.
b. *Jane earns not much money.
c. *Sally talked to Bob about not many problems.

In these examples, it appears that the sequence *not*^NP may occupy the subject position and no other positions (e.g. the object of a verb and the object of a preposition). Postal(1974) suggests:

(6) *Not*-initial NPs occur only in (derived) subject positions.

Postal(1974) points out that the sequence NP^*alone* also exhibits a similar distributional asymmetry.[3]

(7) a. Gronzmeyer alone can help you.
b. Jones alone knows the secret formula.
c. My uncle alone was able to survive.

(8) a. *Call Bob alone.
b. *I talked to Smith alone about the wombat question.
c. *I refuse to work with her alone.

Combining this observation and (6) together, the following seems to hold in English.

(9) *Not*-initial NPs and *alone*-final NPs can occur only in (derived) subject positions.

These phenomena raise a very interesting question for the theory of grammar: what component of grammar can optimally handle such an asymmetry in syntactic distribution? In other words, are elements of grammar currently presupposed by linguists working in the framework of autonomous syntax sufficient? We will argue in the rest of the paper that the notion of categories must be enriched to the effect that a phrase belongs to a syntactic category as well as to some semantic category. We will show that semantically-oriented categorization can provide a natural explanation for the phenomena we are now considering. A couple of preliminary remarks are in order before we present the details of our analysis.

To be strict, (9) is not even descriptively adequate. There are two kinds of systematic exceptions to (9). First, *not* and *alone* can modify certain adverbial phrases fairly freely.

(10) a. Not many years before that, I met him in London.
b. I met him in London not many years before that.

(11) a. In 1984 alone, more than a million people died from cancer.
b. More than a million people died from cancer in 1984 alone.

These are kinds of sentence adverbs that may be positioned sentence-initially as well sentence-finally.

Second, when a phrase is topicalized, both *not* and *alone* seem to be able to modify it.

(12) a. Not many girls would Jack dance with.
b. Not often did I attend physics class.

(13) a. In that way alone can we be sure of winning.
b. These girls alone can we trust at this time.

Note that these constructions are different from the sentence adverb cases in (10) and (11). (12) and (13) become unacceptable if the subject-Aux inversion fails to apply or if the preposing does not happen at all.

(14) a. *Not many girls Jack would dance with.
b. *Not often I attended physics class.

(15) a. *In that way alone we can be sure of winning.
b. *These girls alone we can trust at this time.

(16) a. *Jack would dance with not many girls.
b. *I attended physics class not often.

(16) a. *We can be sure of winning in that way alone.
b. *We can trust these girls alone at this time.

Furthermore, since the inversion is a root phenomenon according to Emonds(1972), (12) and (13) are not acceptable as complement sentences.

(18) a. *I think (that) not many girls would Jack dance with.
b. *They know (that) not often did I attend physics class.

(19) a. *I am told (that) in that way alone can we be sure of winning.
b. *It seems (that) these girls alone can we trust at this time.

It appears, therefore, that (9) is a correct generalization as far as embedded sentences are concerned, maybe with sentence adverbs being the only exception.[4]

2. The Two Analyses

There are at least two potentially promising analyses of *not*-initial NPs. Lasnik(1972), for instance, compares two analyses which he calls the pre-S and the determiner analyses. The pre-S analysis generates *not* in the Comp position.

(20) $[_{S'}$ not $[_{S}$ many people showed up]]

The determiner analysis, on the contrary, assumes that *not many people* in the same sentence is a constituent of type NP.

(21) $[_{S}$ $[_{NP}$ not many people] showed up]

These two analyses, however, have different types of problems and neither of them is fully adequate.

First, let us examine what inadequacies are associated with the pre-S analysis. Of course, we have to recognize that the pre-S has one nice feature: it predicts correctly that *not*-initial NPs do not appear in object positions. Nevertheless, this analysis has an obvious problem. Note that there is a strong cooccurrence restriction holding between *not* and the NP that follows it. Consider the following paradigm.

(22) a. Not {every unicorn / many unicorns / all unicorns} can be captured alive.
b. Not much water is left.

(23) a. *Not {some unicorns / most unicorns / the unicorn / each unicorn} can be captured alive.
b. *Not Bill came.

The pre-S analysis, in principle, should not discriminate the sentences in

(22) from those in (23). It is very unlikely that there is any selectional restriction holding between an element in Comp and the subject NP. On the other hand, this restriction can most naturally be stated if we analyze the construction in question along the determiner analysis.[5] Thus, the pre-S analysis must be rejected on the ground that it will generate the sentences in (20) without any ad hoc and unnatural stipulation.

A second problem associated with the pre-S analysis is that it cannot be extended in any direction so as to cover *alone*-final NPs since *alone* obviously appears in a postnominal position. Hence, if *not*-initial NPs and *alone*-final NPs are two realizations of the same single phenomenon, the pre-S analysis will miss a significant generalization.

Finally, if coordination gives us any clue regarding the phrase structure, in particular the constituenthood, the following data suggest that *not* ^NP is actually a constituent.

(24) a. Not many colonels and not many majors were demoted yesterday.
b. Not much wheat and not much barley was sold to the Turks.
(Postal(1974))

In fact, many native speakers have the intuition that *not*-initial NPs are constituents. Lasnik(1972), for example, reports the same intuition.

(25) "It is my intuition that strings such as *not many men, not often* are surface structure constituents. I have no conclusive syntactic arguments that this is the case, but I will assume that they are constituents by some stage in the derivation." (Lasnik(1972:26))

Note incidentally that the critical remarks given above apply as they are to any one of the family of analyses that can be termed "a pre-S analysis". Thus, as far as *not* and the following NP do not form a constituent, the following analyses must all be rejected on the same grounds.

(26) a. [$_S$ not [$_S$ many people showed up]]
b. [$_S$ not many people showed up]

The determiner analysis (or the family of equivalent analyses) may overcome the problems discussed so far and yet it embodies a fatal flaw. The distribution of *not*-initial NPs and *alone*-final NPs remain totally unexplained. If *not*-initial NPs and *alone*-final NPs are really constituents, and furthermore if they are NPs, then they should be able to occur under any NP positions allowed by the grammar, for instance, in object positions as well as in subject positions.

This is a very paradoxical situation. One analysis captures some properties but leaves others unexplained. The other analysis captures the properties that are left unexplained in the first analysis but loses explanation for the properties that are taken care of by the first analysis. They complement each other but neither of them is fully adequate.

3. Semantic Categories

I would like to argue that a sentence is well-formed if and only if it is syntactically well-formed and semantically coherent. Here, the notion of semantic coherence is used in a special way that probably most closely resembles the use of the term "coherence" by Ajdukiewicz. Each lexical item not only belongs to some syntactic category but also to some semantic category or s-category. Cf. Ajdukiewicz(1935), Bach(1981), Bar-Hillel(1953), and Lambek(1961). Let us say we have two primitive categories, *t* and *e*. *t* is the category of sentences (they have *t*ruth values) and *e* is the category of individuals (or *e*ntities). On the basis of these two primitive categories, we define an infinite set of derived categories.

(27) i. t and e are s-categories.
ii. If α and β are s-categories, α/β is also an s-category.
iii. Nothing else is an s-category.

S-categories of the form α/β are often called functor categories because they denote a category of functions whose domain consists of elements in category β and whose range consists of elements in category α.

On the basis of these s-categories, the following is defined as an axiom of our theory.

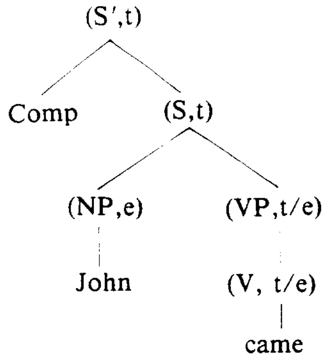
(28) A node X is s-coherent in s-category α if its two daughter nodes Y and Z belong to s-categories α/β and β, respectively.

This presupposes that every syntactic node is at most binary branching. Let us label each node in a syntactic tree by a pair (K,L), where K is a syntactic category and L is a s-category, if the node is s-coherent in L. We further assume the following.

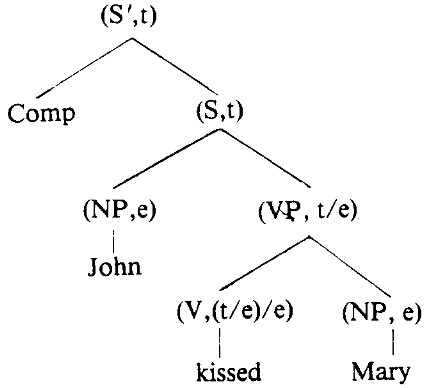
(29) A syntactic tree is well-formed if and only if the root is s-coherent in t.

The following are some typical examples of well-formed syntactic representations.

(30)



(31)



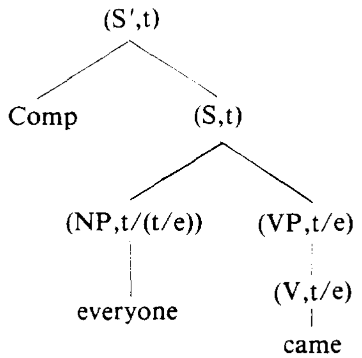
Terminal elements like *Mary* and *came* are s-coherent by definition. Their s-categories are specified in the lexicon. Note that the Comp node in (30) and (31), which does not dominate any terminal string, acts as an identity function with respect to s-categories. Thus, for any node x with the pair (S,y), y being some s-category, the node immediately and exclusively dominating x and Comp that is empty will have the pair (S′, y). A non-branching node is treated on a par; we assume that a non-branching node dominates an empty category which is an identity function as well as its single daughter. By this convention, (VP,t/e) in (30) proves to be s-coherent.

In (30) and (31), we analyzed NPs as e's, and accordingly other categories. This can be summarized as follows.

(32) NP: e
VP: t/e transitive verb: (t/e)/e

This requires some revision. If we consider only proper names like *John* and *Mary,* we can maintain a simple system like the one we have just outlined. But once the scope of investigation is expanded to include quantificational phrases such as *everyone* and *no one,* we cannot maintain the system as it is. Since the quantificational pharases are not referential expressions, they cannot be regarded as individuals. Rather, we take them as a set of properties. Cf. Montague(1974a). Now, all quantificational phrases belong to s-category t/(t/e).

(33)



Following Montague(1974c), we analyze proper names on a par with quantifiers. Proper names are also analyzed as a set of properties that an individual has.

(34) *everyone* $\rightarrow \lambda P \forall x_1[\text{person}'(x_1) \rightarrow P(x_1)]$[6]
John $\rightarrow \lambda P[P(j)]$
$[_{NP}\ e_i]$ $\rightarrow \lambda P[P(x_i)]$

In the above, I have included the translation for an empty NP that is coindexed with some operator and functions as a variable.

We also modify the category of VPs. We assume that NPs are generally arguments and VPs and transitive verbs are functions that take NPs are arguments. Cf. Keenan and Faltz(1978) and Montague(1974b).[7] Using the phrase structure rules presented in (2), we will analyze the tense as functions on VPs along the line suggested by Bach(1980). Thus, we take VPs as elements of t/e and VP's as elements of t/(t/(t/e)).

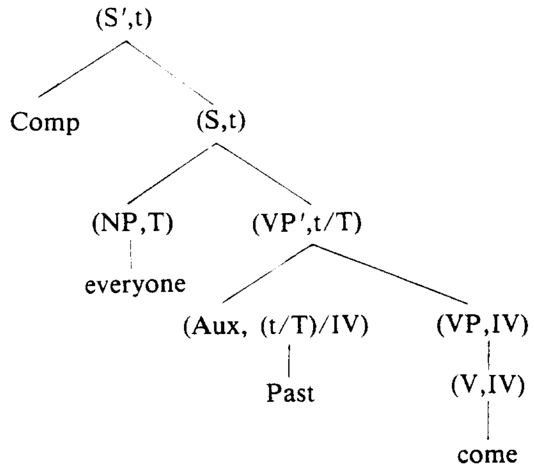
(35) NP: t/(t/e)
VP: t/e
Aux: (t/(t/(t/e)))/(t/e)
VP': t/(t/(t/e))

To improve the readability we use abbreviations:

(36) NP: T = t/IV = t/(t/e)
VP: IV = t/e
Aux: (t/T)/IV
VP′: t/T

Now (33) is reanalyzed as:

(37)



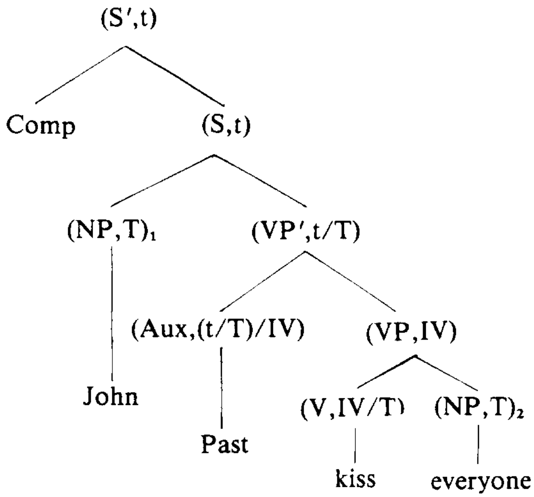
A well-formed syntactic tree of this kind is compositionally translated into logical expressions. That is, each synatactic node that is s-coherent in α will have the corresponding logical expression of that type. Note that we can give a unique translation to each node. If a node that belongs to α/β translates into X and a node that belongs to β translates into Y, then the node dominating these two nodes, which belongs to α, will translate into X(Y), the semantic counterpart of function-argument application. Thus the tree in (37) will involve the following translations. Again, empty categories other than coindexed traces have no semantic contributions.

(38) (V,IV) → come′
(VP,IV) → come′
(Aux,(t/T)/IV) → $\lambda P \lambda X[HX(P)]$
(VP′,t/T) → $\lambda P \lambda X[HX(P)](come')$
$= \lambda X[HX(come')]$
(NP,T) → $\lambda P \forall x[person'(x) \rightarrow P(x)]$

$$
\begin{aligned}
(S,t) &\rightarrow \lambda X[HX(come')](\lambda P\forall x[person'(x) \rightarrow P(x)]) \\
&= H\lambda P\forall x[person'(x) \rightarrow P(x)](come') \\
&= H\forall x[person'(x) \rightarrow come'(x)] \\
(S',t) &\rightarrow H\forall x[person'(x) \rightarrow come'(x)]
\end{aligned}
$$

The following is an example that involves a transitive verb.

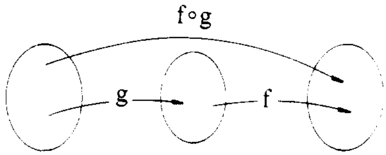
(39)



(40)
$$
\begin{aligned}
(NP,T)_2 &\rightarrow \lambda P\forall y[person'(y) \rightarrow P(y)] \\
(V,IV/T) &\rightarrow \lambda X\lambda z[X(\lambda x[kiss'(x)\ (z)])] \\
(VP,IV) &\rightarrow \lambda X\lambda z[X(\lambda x[kiss'(x)\ (z)])]\ (\lambda P\forall y[person'(y) \rightarrow P(y)]) \\
&= \lambda z[\lambda P\forall y[person'(y) \rightarrow P(y)](\lambda x[kiss'(x)\ (z)])] \\
&= \lambda z[\forall y[person'(y) \rightarrow \lambda x[kiss'(x)\ (z)](y)]] \\
&= \lambda z[\forall y[person'(y) \rightarrow kiss'(y)(z)]] \\
(Aux,\ (t/T)/IV) &\rightarrow \lambda P\lambda X[HX(P)] \\
(VP',t/T) &\rightarrow \lambda P\lambda X[HX(P)](\lambda z\forall y[person'(y) \rightarrow kiss'(y)(z)]) \\
&= \lambda X[HX(\lambda z\forall y[person'(y) \rightarrow kiss'(y)(z)])] \\
(NP,T)_1 &\rightarrow \lambda PP(j) \\
(S,t) &\rightarrow \lambda X[HX(\lambda z\forall y[person'(y) \rightarrow kiss'(y)(z)])](\lambda PP(j)) \\
&= H\lambda PP(j)\ (\lambda z\forall y[person'(y) \rightarrow kiss'(y)(z)])]) \\
&= H\lambda z\forall y[person'(y) \rightarrow kiss'(y)(z)](j) \\
&= H\forall y[person'(y) \rightarrow kiss'(y)(j)] \\
(S',t) &\rightarrow H\forall y[person'(y) \rightarrow kiss'(y)(j)]
\end{aligned}
$$

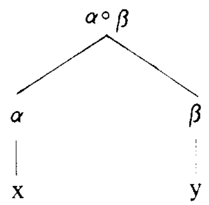
Our theory of s-categories is supplemented by function composition. Function composition allows us to combine two functor categories when the domain of one function coincides with the range of another function.

(41)



By the definition of function composition, the translation for such a composite category is as follows.

(42) Function Composition (FC):



If x translates into x′ and y into y′, then ($\alpha \circ \beta$) translates into $\lambda u\,[x'\,(y'\,(u))]$.

Since FC combines two functions, α and β in (42) must be of functor categories. In our categorial system, this means that the node dominating x and y is s-coherent in some composite function category if x belongs to α/δ and y belongs to δ/β.

(43) A node X is s-coherent in s-category $\alpha/\delta \circ \delta/\beta$ if its two daughter nodes Y and Z belong to s-categories α/δ and δ/β, respectively.[8]

By combining (43) and the previous definition of s-coherence in (28), we can define the general notion of s-coherence as follows.

(44) A node X is s-cohrent if and only if X is s-coherent in some s-category according to either (28) or (43).

We will use FC only for a specific purpose. Cf. Abe(1984) for the use of FC in a totally different area. That is, to ensure the s-coherence of the *not*-initial NP and the *alone*-final NP, we analyze them as belonging to some composite categories. As we saw in section 2, we must alanlyze these as constituents to state the necessary cooccurrence restriction in a natural fashion. *not* and *alone* can most naturally be fit in t/t and IV/IV, respectively. Observe the following.

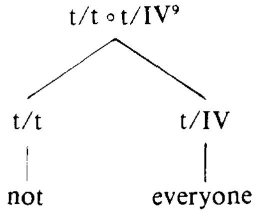
(45) Not that I hate fish, but ...

(46) John came back alone.

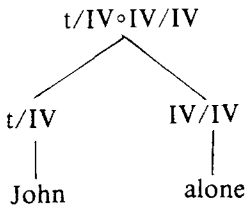
The following phrases become s-coherent only through FC.

(47) a. not everyone
b. John alone

(48)



(49)



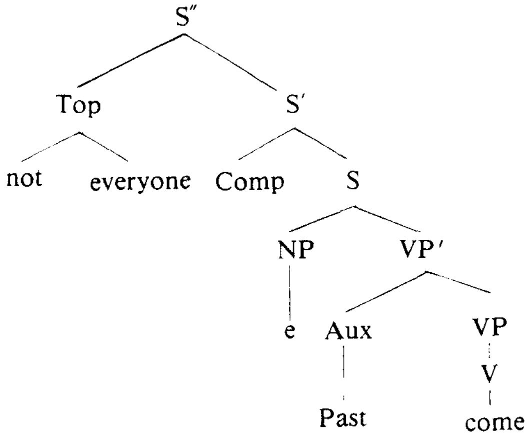
Given this much of mechanism, let us present an analysis of *not*-initial NPs and *alone*-final NPs which will overcome the problems discussed in Section 2 in a revealing fashion.

4. Topic Construction

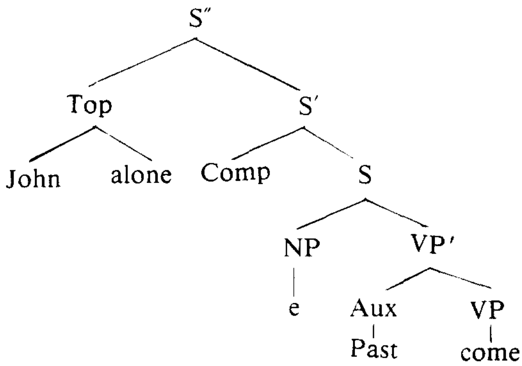
Our analysis is a hybrid of the two analyses discussed in Section 2, the pre-S analysis and the determiner analysis. We will maintain that *not*-initial NPs and *alone*-final NPs are constituents so that the appropriate selectional restictions can be stated locally. We will assume, however, that *not*-initial NPs and *alone*-final NPs are not in normal argument positions but that they are in topic positions. The following sentences will thus have the syntactic representations given in (51), where s-categories are suppressed for convenience.

(50) a. Not everyone came.
b. John alone came.

(51) a.



b.



This presupposes the following phrase structure rules that are proposed in Chomsky(1977).

(52) $S'' \rightarrow \text{Top } S'$

$$S' \rightarrow \text{Comp} \begin{Bmatrix} S'' \\ S \end{Bmatrix}$$

Furthermore, we assume that there is a movement of the empty element inside S to the Comp position. The empty element in Comp functions as the operator binding the empty trace in S.

(53) a. [not everyone [e_i [e_i Past come]]]
b. [John alone [e_i [e_i Past come]]]

The empty operator is needed so that the topic phrase can be correctly related to the rest of the sentence through the rule of "predication" (cf. Chomsky(1977, 1981, 1982)), the exact nature of which will be discussed later.

This analysis nicely captures the fact that these phrases are not allowed in object positions.

(54) a. *John invited not everyone.
b. *John invited Mary alone.

It also correctly predicts that the sentences in (54) have the grammatical counterparts, in which the *not*-initial NP and the *alone*-final NP are generated in the topic position.

(55) a. Not everyone did John invite.
b. Mary alone did John invite

(56) a. [not everyone [e_i [John Past invite e_i]]]
b. [Mary alone [e_i [John Past invite e_i]]]

Here, we assume that the tense will be placed in Comp later in the Phonetic Component (PF) in the sense of Chomsky(1981), so that the apparent inversion effect is achieved.

We have noted earlier that there is an apparent asymmetry between the matrix and the complement clauses with respect to the distribution of *not*-initial NPs and *alone*-final NPs.

(57) a. I think not everyone came.
b. I think Mary alone came.

(58) a. *I think not everyone did John invite.
b. *I think Mary alone did John invite.

This can be attributed to the fact that the inversion is triggered only when the topic phrase corresponds to nonsubjects. Thus, the inversion is not necessary when the topic corresponds to the subject as in (57), whereas (58) requires the inversion because the topic corresponds to the object. However, the inversion is apparently restricted to matrix clauses, a restriction that applies in the PF component. Hence, the structures in (58) are ruled out in the PF component.

Note that S-internal adverbs can be treated in a parallel way. It is well-formed in the topic position of the matrix clause but nowhere else.

(59) a. Not often did John come.
b. [not often [e_i [John Past come e_i]]]

(60) a. *John came not often.
b. *I think not often did John come.

This is in direct contrast to sentence adverbs, which are a lot freer.

(61) a. Not many years before that, John met Mary.
b. John met Mary, not many years before that.
c. It is true that not many years before that John met Mary.
d. It is true that John met Mary not many years before that.

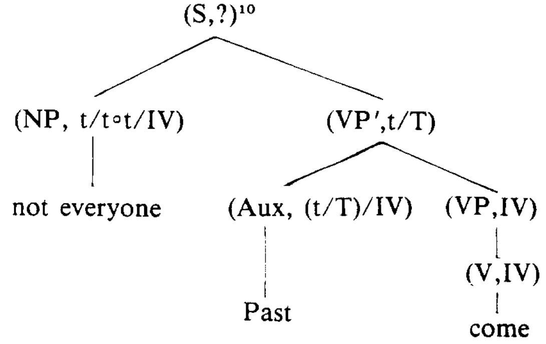
I propose that these adverbial phrases are sentence modifiers and that they are Chomsky-adjoined to S.

(62) [$_S$ not many years before that [$_S$ John Past meet Mary]]

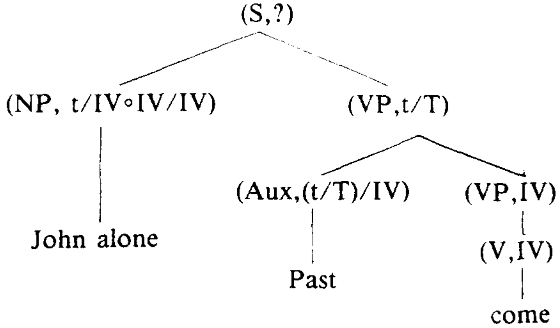
Since these cases do not involve topicalization, there is no binding of empty categories or the inversion of Tense.

Let us now look at the facts more closely and seek a formal explanation as to why *not*-initial NPs and *alone*-final NPs are only well-formed in the topic position. The reason why these phrases do not appear inside S or in A(rgument)-positions in the sense of Chomsky(1981) can be stated in terms of s-categories.

(63)



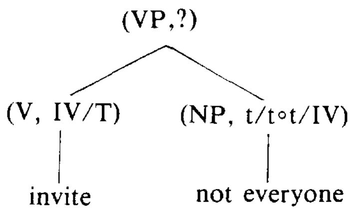
(64)



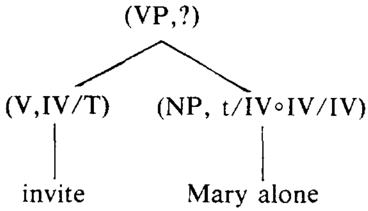
Being composite functions of type t/t∘t/IV and t/IV∘IV/IV, respectively, *not*-initial NPs and *alone*-final NPs cannot occupy subject positions. The domain of composite functions of type t/t∘t/IV is a set of entities of type IV, and hence *not everyone* in (63) cannot take VP′ *Past come* as its argument since the latter is of type t/T. Similarly, the domain of composite function of type t/IV∘IV/IV is a set of entities of type IV, and therefore *John alone* cannot take VP′ *Past come* as its argument. Consequently, the root S(or S′)can never be taken as s-coherent, violating the well-formedness condition (29).

The same explanation carries over to the object case. Here, we simply provide a part of the entire tree since we can easily detect the lack of s-coherence in this local domain.

(65)



(66)



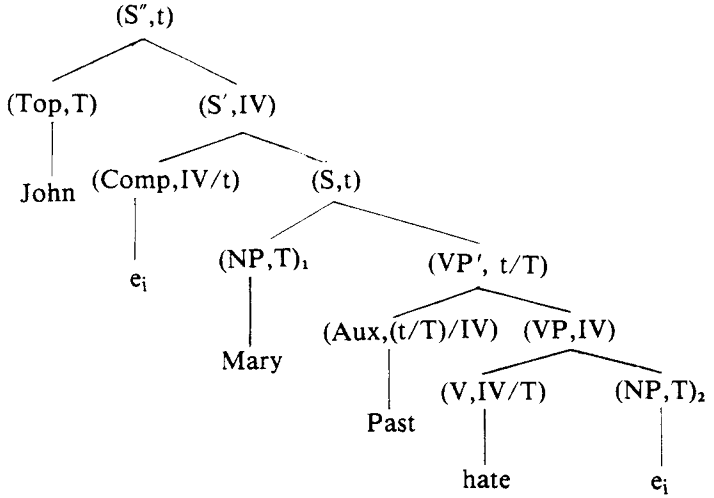
Here, the object NP is not of appropriate s-categories so that the verb *invite* cannot take in its argument.[11]

Let us now demonstrate how the topic structure allows *not*-initial NPs and *alone*-final NPs. This proceeds in two steps. First, I will show that the relevant structure is in fact s-coherent. Then, I will show that the correct interpretation can be obtained from the proposed configuration.

The following is a case of simple topicalization.

(67) John, Mary hated.

(68)



The crucial part of (68) is the s-category assigned to the empty category in Comp. This empty element, which is coindexed with the trace in the object position, functions as the binder of the trace. It turns the sentence containing a free (in S) variable into a predicate, which will then be predicated of the topic phrase. Hence, the empty binder is of s-category IV/t. Now the topic, which is of T (= t/IV), is a function taking this derived predicate as an argument.

The translation procedures up to (S,t) involve nothing new.

(69) $(NP,T)_2 \rightarrow \lambda PP(x_i)$
$(V,IV/T) \rightarrow \lambda X \lambda y[X(\lambda x[hate'(x)(y)])]$
$(VP,IV) \rightarrow \lambda X \lambda y[X(\lambda x[hate'(x)(y)])](\lambda PP(x_i))$
$\rightarrow \lambda y[\lambda PP(x_i)(\lambda x[hate'(x)(y)])]$
$\rightarrow \lambda y[\lambda x[hate'(x)(y)](x_i)]$
$\rightarrow \lambda y\ [hate'(x_i)(y)]$

$$(Aux,(t/T)/IV) \rightarrow \lambda P\lambda X[HX(P)]$$
$$(VP',t/T) \rightarrow \lambda P\lambda X[HX(P)](\lambda y[hate'(x_i)(y)])$$
$$= \lambda X[HX(\lambda y[hate'(x_i)(y)])]$$
$$(NP,T)_i \rightarrow \lambda PP(m)$$
$$(S,t) \rightarrow \lambda X[HX(\lambda y[hate'(x_i)(y)])](\lambda PP(m))$$
$$\rightarrow H\lambda PP(m)(\lambda y[hate'(x_i)(y)])$$
$$\rightarrow H\lambda y[hate'(x_i)(y)](m)$$
$$\rightarrow H\ hate'(x_i)(m)$$

Thus, the substring *Mary Past hate* e_i receives the translation $Hhate'(x_i)(m)$, which is an open sentence. The translation for (S', IV) is obtained through the following procedural assignment, rather than by s-categories.

(70) Given the configuration $[_X(Comp,t/e), Y]$, X translates into $\lambda x_i\phi$, where (Comp,t/e) dominates e_i and Y translates into ϕ.

This simply is a formalization of the notion "the empty operator binds the trace". Note that (Comp,t/e) arises only when it contains the empty element that is coindexed with another empty element in some A-position. Now the rest proceeds as follows.

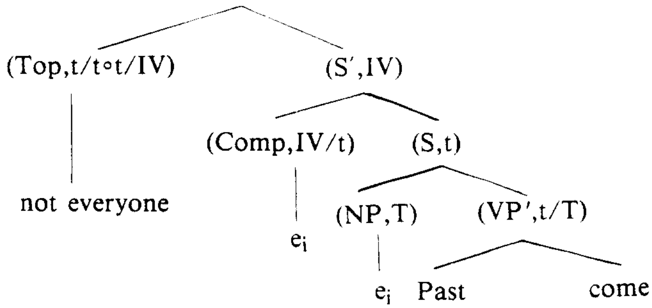
(71) $(S',IV) \rightarrow \lambda x_i[Hhate'(x_i)(m)]$ by (70)
$(Top,T) \rightarrow \lambda PP(j)$
$(S'',t) \rightarrow \lambda PP(j)(\lambda x_i(Hhate'(x_i)(m)])$
$= \lambda x_i[Hhate'(x_i)(m)](j)$
$= Hhate'(j)(m)$

The logical expression given to the whole sentence in (68) is identical to the one that would be given to the non-topicalized version, *Mary hated John.*

Given this, it is easy to show that the following sentences are all s-coherent and that they receive desired translations.

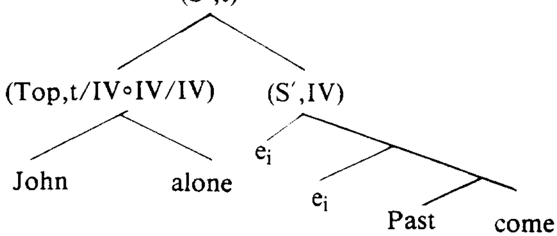
(72) a. Not everyone came.
b. John alone came.
c. Not everyone did John invite.
d. Mary alone did John invite.

(73)



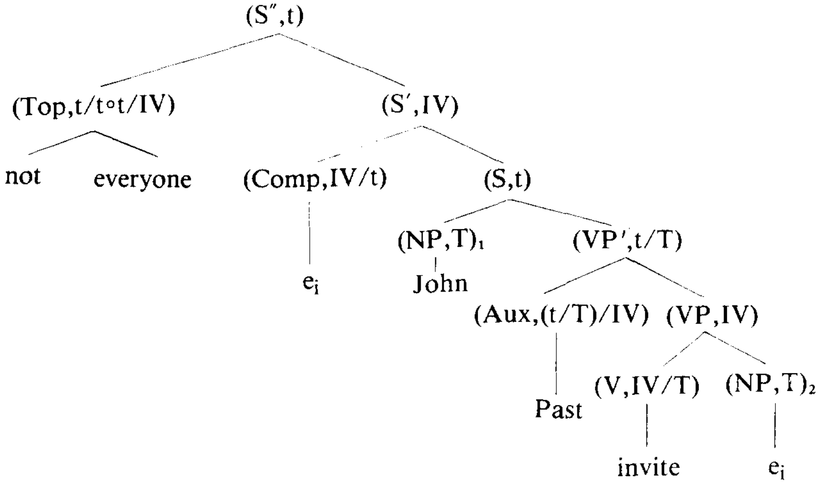
(74)
$$
\begin{aligned}
&(VP',t/T) \rightarrow \lambda P \lambda X[HX(P)](come') = \lambda X[HX(come')] \\
&(NP,T) \rightarrow \lambda P P(x_i) \\
&(S,t) \rightarrow \lambda X[HX(come')](\lambda P P(x_i)) \\
&\qquad = H\lambda P P(x_i)(come') = Hcome'(x_i) \\
&(S',IV) \rightarrow \lambda x_i[Hcome'(x_i)] \\
&(Top,t/t\circ t/IV) \rightarrow \lambda P[\sim(\lambda Q \forall x[person'(x) \rightarrow Q(x)](P))] \\
&\qquad\qquad = \lambda P[\sim \forall x[person'(x) \rightarrow P(x)]] \\
&(S'',t) \rightarrow \lambda P[\sim \forall x[person'(x) \rightarrow P(x)]](\lambda x_i[Hcome'(x_i])) \\
&\qquad = \sim \forall x[person'(x) \rightarrow \lambda x_i[Hcome'(x_i)](x)] \\
&\qquad = \sim \forall x[person'(x) \rightarrow Hcome'(x)]
\end{aligned}
$$

(75)



(76)
$$
\begin{aligned}
&(S',IV) \rightarrow \lambda x_i[Hcome'(x_i)] \\
&(Top, t/IV\circ IV/IV) \rightarrow \lambda Q[\lambda P P(j)(\lambda R \lambda y \forall x[P(x) \leftrightarrow x=y](Q))] \\
&\qquad = \lambda Q[\lambda P P(j)(\lambda y \forall x[Q(x) \leftrightarrow x=y])] \\
&\qquad = \lambda Q[\lambda y \forall x[Q(x) \leftrightarrow x=y](j)] \\
&\qquad \doteq \lambda Q \forall x[Q(x) \leftrightarrow x = y] \\
&(S'',t) \rightarrow \lambda Q \forall x[Q(x) \leftrightarrow x=j](\lambda x_i[Hcome'(x_i)]) \\
&\qquad = \forall x[\lambda x_i[Hcome'(x_i)](x) \leftrightarrow x=j] \\
&\qquad = \forall x[Hcome'(x_i) \leftrightarrow x=j]
\end{aligned}
$$

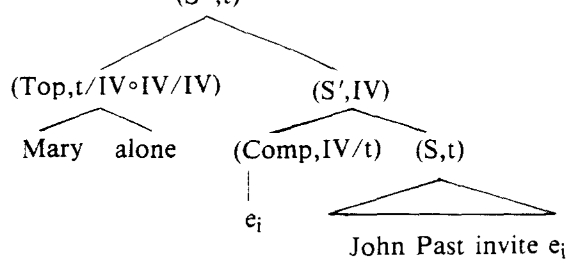
(77)



(78)

$$
\begin{aligned}
&(NP,T)_2 \rightarrow \lambda P P(x_i)\\
&(V,IV/T) \rightarrow \lambda X \lambda y[X(\lambda x[\text{invite}'(x)(y)])]\\
&(VP,IV) \rightarrow \lambda X \lambda y[X(\lambda x[\text{invite}'(x)(y)])](\lambda P P(x_i))\\
&\qquad = \lambda y[\lambda P P(x_i)(\lambda x[\text{invite}'(x)(y)])]\\
&\qquad = \lambda y[\lambda x[\text{invite}'(x)(y)](x_i)]\\
&\qquad = \lambda y[\text{invite}'(x_i)(y)]\\
&(Aux,(t/T)/IV) \rightarrow \lambda P \lambda X[HX(P)]\\
&(VP',t/T) \rightarrow \lambda P \lambda X[HX(P)](\lambda y[\text{invite}'(x_i)(y)])\\
&\qquad = \lambda X[HX(\lambda y[\text{invite}'(x_i)(y)])]\\
&(NP,T)_1 \rightarrow \lambda P P(j)\\
&(S,t) \rightarrow \lambda X[HX(\lambda y[\text{invite}'(x_i)(y)])](\lambda P P(j))\\
&\qquad = H\lambda P P(j)(\lambda y[\text{invite}'(x_i)(y)])\\
&\qquad = H\lambda y[\text{invite}'(x_i)(y)](j)\\
&\qquad = H\text{invite}'(x_i)(j)\\
&(S',IV) \rightarrow \lambda x_i[H\text{invite}'(x_i)(j)]\\
&(Top,t/t\circ t/IV) \rightarrow \lambda P[\sim\forall x[\text{person}'(x) \rightarrow P(x)]]\\
&(S'',t) \rightarrow \lambda P[\sim\forall x[\text{person}'(x) \rightarrow P(x)]](\lambda x_i[H\text{invite}'(x_i)(j)])\\
&\qquad = \sim\forall x[\text{person}'(x) \rightarrow \lambda x_i[H\text{invite}'(x_i)(j)](x)]\\
&\qquad = \sim\forall x[\text{person}'(x) \rightarrow H\text{invite}'(x)(j)]
\end{aligned}
$$

(79)



(80) $(S, t) \rightarrow \text{Hinvite}'(x_i)(j)$
$(S', IV) \rightarrow \lambda x_i[\text{Hinvite}'(x_i)(j)]$
$(Top, t/IV \circ IV/IV) \rightarrow \lambda Q \forall x[Q(x) \leftrightarrow x = m]$
$(S'', t) \rightarrow \lambda Q \forall x[Q(x) \leftrightarrow x = m](\lambda x_i[\text{Hinvite}'(x_i)(j)])$
$= \forall x[\lambda x_i(\text{Hinvite}'(x_i)(j)](x) \leftrightarrow x = m]$
$= \forall x[\text{Hinvite}'(x)(j) \leftrightarrow x = m]$

5. Conclusion

We have seen that by generating *not*-initial NPs and *alone*-final NPs in the topic position, all the interesting properties of these phrases can naturally be accounted for. At the beginning, it seemed that the distribution of these phrases is very peculiar. The apparent subject-object asymmetry is, however, attributed to the lack of the inversion of tense when the subject is topicalized. Our assumption is that *not*-initial NPs and *alone*-final NPs are never well-formed in A-positions, and that they are limited to an $\bar{A}$-position, the topic. These assumptions turned out to follow from our theory of s-coherence and the independently necessary theory of topicalization. This paper has demonstrated that a certain distributional property can only be captured through the amalgamation of a syntactic framework that embodies a theory of empty categories and a theory of compositional semantics that, in turn, is based on the theory of s-categories or some of its equivalents.

Footnotes

* This is a revised and extended version of the paper that was read at the Seoul Workshop on Formal Grammar, December 1983. I would like to thank the participants in the workshop for many interesting comments. David Dowty and Kiyong Lee gave me useful comments, which motivated and directed some of the crucial revisions made in this paper.

1. Of course, a category can be subcategorized as in the case of verbs. Verbs

are subcategorized according to the number and the type of complements they take. Thus, it is possible that a certain element in a lexical category has a degenerate distribution in the maximally possible patterns specified by the phrase structure rules. But this does not apply to phrasal categories like NP and PP.

2. In the framework developed by Chomsky(1981), the controlled subject or PRO is restricted to the subject position and this fact is shown to follow from some of the other modules of the theory.

3. *alone* in all of the relevant examples must be interpreted in the same way as *only* is interpreted. Thus, some of the sentences in (8) become well-formed with a different interpretation.

4. Note that we should not treat the following examples on a par with the other examples of *not*-initial NPs discussed in the text.

(i) I invited Jóhn, not Bíll.

(ii) Máry invited Jóhn, not Súsan, Bíll.

These sentences do not share the properties characteristic to the "normal" *not*-initial NP construction.

5. See Barwise and Cooper(1981) for an attempt to explain the cooccurrence restriction in terms of the semantic characteristics of quantifiers.

6. We will be working in a purely extensional model. The following table summarizes the convention used in this paper.

	constants	*variables*	*category*
individuals	j, m,...	$x,y,z,x_i,\ldots$	e
properties	walk′, run′,..	P,Q,R,...	t/e
sets of properties		X,Y, ...	t/(t/e)

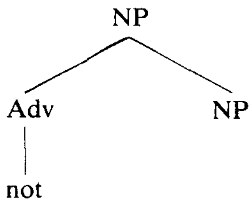
7. This position is in direct contrast to the PTQ type analysis of subject and predicate, on which my earlier analysis of *not*-initial NPs is based. Cf. Abe (1983).

8. In order to include categories of composite functions, an obvious revision is necessary in the recursive difinition of (27). The following clause must be addded.

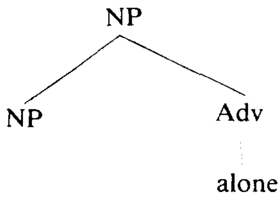
(i) If α/δ and δ/β are s-categories, $\alpha/\delta \circ \delta/\beta$ is also an s-category.

9. Syntactically, *not* and *alone* can probablly be analyzed as external modifiers that result in purely endocentric constructions.

(i)

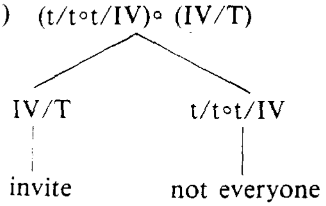


(ii)

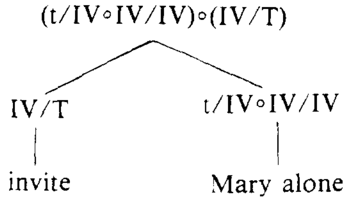


10. For the internal structure of *not*-initial NPs and *alone*-final NPs, see (48), (49) and Footnote 9.
11. We must note that in (65) and (66), it is logically possible to form a composite function in the following way.

(i)



(ii)



This possibility is easily ruled out. Such VPs cannot combine properly with the tense to form VP′. Since the tense is of s-category (t/T)/IV, VP′ would not be s-coherent even if those complicated composite functions in (i) and (ii) were actually formed.

References

Abe, Y. (1983) "Montague Grammar and Negation" in *Issues in Syntax and Semantics: Festschrift for Masatake Muraki,* 1-11, Sansyusya Co., Tokyo.

________ (1985) *A Theory of Categorial Morphology and Agglutination in Japanese*, unpublished Ph. D. dissertation, University of Massachusetts at Amherst.

Ajdukiewicz, K. (1935) "Die syntaktische konnexität" *Studia Philosophica* 1, 1-27. (Appears in English translation in Storrs McCall, ed., *Polish Logic,* Oxford Univ. Press)

Bach, E. (1980) "Tenses and aspects as functions on verb-phrases," in *Time, Tense, and Quantifiers: Proceedings of the Stuttgart Conference on the Logic of Tense and Quantification,* ed. by C. Rohrer, pp. 19-37, Max Niemeyer Verlag.

________ (1981) "Generalized Categorial Grammars and the English Auxiliary," ms., Univ. of Massachusetts, Amherst.

Bar-Hillel, Y. (1953) "A Quasi-arithmatical Notation for Syntactic Description," *Language* 29.1, 49-58.

Barwise, J. and R. Cooper (1981) "Generalized Quantifiers and Natural Languages," *Linguistics and Philosophy* 4, 159-319.

Chomsky, N. (1977) "On *Wh*-movement," in *Formal Syntax,* ed. by P. Culicover, T. Wasow, and A. Akmajian, 71-132, Academic Press, New York.

_______ (1981) *Lectures on Government and Binding,* Foris, Dordrecht.

_______ (1982) *Some Concepts and Consequences of the Theory of Government and Binding,* MIT Press.

Emonds, J. (1976) *A Transformational Approach to English Syntax: Root, Structure-Preserving, and Local Transformations,* Academic Press, New York.

Keenan, E.L. and L.M. Faltz (1978) "Logical Types for Natural Language," *UCLA Occasional Papers in Linguistics 3.*

Klima, E.S. (1964) "Negation in English," in *The Structure of Language,* ed. by J. Fodor and J. Katz, 246-323, Prentice-Hall, Englewood Cliffs, NJ.

Lambek, J. (1961) "On the Calculus of Syntactic Types," in *Structure of Language and Its Mathematical Aspects: Proceedings of the Twelfth Symposium in Applied Mathematics,* ed. by R. Jakobson, 166-178, American Mathematical Society.

Lasnik, H. (1972) *Analyses of Negation in English*, unpublished Ph. D dissertation MIT.

Montague, R. (1974a) *Formal Philosophy: Selected Papers of Richard Montague*, Yale University Press.

_______ (1974b) "English as a Formal Language," in Montague (1974a) 188-221.

_______ (1974c) "The Proper Treatment of Quantification in Ordinary English," in Montague (1974a), 247-270.

Partee, B. (1970) "Negation, Conjunction, and Quantifiers: Syntax vs. Semantics," *Foundations of Language* 6.2, 153-165.

Postal, P. (1974) *On Raising,* MIT Press.

Rikkyo University
Tokyo, Japan