On the Viability of the Notion of 'Subject' in Universal Grammar

William A. Foley

and Robert D. Van Valin, Jr.

Australian National University. University of California, Berkeley.

1. Since Plato the notion of subject has figured prominently in grammatical theory. During the era of transformational grammar this notion lost its theoretical and descriptive significance because of the re-definition of grammatical relations in terms of constituent structure. Recently, however, the notion of subject has again become an issue in linguistic theory. This renewed interest was sparked by Keenan and Comrie's (1972) landmark study of cross-linguistic variation in relativization strategies and by Postal and Perlmutter's (1974) relationally based theory of grammar. Keenan and Comrie (1972) argue that notions like subject, direct object and indirect object are necessary to account for their cross-linguistic generalizations. Postal and Perlmutter (1974) attempt to recast the findings of transformational grammar within their theory of relational grammar, taking the notion of subject as a primitive.

Obviously, if these theoretical generalizations are to have any validity, it is necessary to give a universal characterization of subjecthood. Therefore, one must provide a methodology by which one can identify that noun phrase which functions as a subject within the grammatical system of a language. Keenan (1976) is an attempt to do this for the 'basic' sentences in a language. He proposes 30-odd properties which he claims can be used to identify the subject noun phrase, if any, in a basic sentence in any given language. None of the properties, however, can be shown to be necessary or sufficient for defining subject- hood. Keenan nevertheless maintains that subject is a universal notion which is accordingly realized in the grammar of every natural language. His task is then to reconcile the fact that there are no universal defining characteristics of 'subject' with the assertion that it is a universal notion. His suggestion is to define subject as a multifactor concept, where "the subjecthood of an NP (in a sentence) is a matter of degree." (Keenan 1976:307), i.e. the more subject properties an NP has, the more subject it is. Therefore, we may conclude from this that the NP in a basic sentence with the fullest complement of subject properties is the subject of that sentence. Furthermore, Keenan claims that the NP thus identified should "in clear cases" agree with our "pretheoretical intuition" about which NP is the subject of that sentence. (Keenan 1976:306) This is a highly problematic claim in that first of all, no criteria are given to define what counts as a "clear case", and second, one's "pretheoretical intuitions" about the subjecthood of an NP are questionable outside of the analyst's native language.

If the subjecthood of an NP in a sentence is a given language is always a matter of degree, then we might expect the identification of subject in some languages to be problematic, i.e. the spread of subject properties among the NP's in a sentence may be so diffuse
as to render the non-ad hoc identification of an NP as subject impossible. Cases of this have been reported for Philippine languages by Schachter (1976) and for Barai (Papua New Guinea) by Olson (1976). Schachter has demonstrated that Keenan's subject properties are split between two NPs in Tagalog (the Actor and the Topic) and he argues that neither NP can be specified as the subject of a Tagalog sentence even in terms of Keenan's multifactor definition. In a later paper (Schachter 1977) he argues that this split organization of subject properties in Tagalog is a reflection of the basic universal organization of subject properties.

2. It is the purpose of this paper to examine the distribution of subject properties in order to see if it is viable as a construct in universal grammar, taking Schachter's (1977) analysis as our starting point. We will examine this distribution in three languages: Tagalog, Navajo and Lakhota. We will show that subject properties do not divide up neatly in all languages between Actor (role) properties and Topic (referential) properties, but that language specific typological facts influence their distribution, so that in order to make implicative statements about the distribution of subject properties in any language, we must investigate factors not considered in Schachter's analysis. The theoretical framework in which this analysis will be presented differs in crucial respects from those competing theories which are presently well known. Accordingly, we shall here sketch out some of the basic assumptions and notions most germane to our present discussion.

We view the basic organization of clause level grammar as the interplay between role functions and referential structures. Role functions represent the basic roles that NPs may play in the clause, and these are usually signalled by case marking and less commonly, word order. The roles are established by the semantics of the verb (Foley 1976a). Referential structure represents the basic organization of the clause in terms of the referentiality of the NPs therein. This referentiality may be established linguistically by either of two considerations. It may be established by discourse controlled factors such as definiteness and givenness (Chafe 1976) or by the inherent referentiality of the NP. There appears to be a universal hierarchy of inherent topic-worthiness called variously the Natural Topic Hierarchy (Hawkinson and Hyman 1975), Inherent Lexical Content Hierarchy (Silverstein 1977) and Referentiality Hierarchy (Foley 1976b). The Hierarchy in universal terms is (Foley 1976b).

speaker > hearer > human proper > human common > animate > inanimate

Referential structure is the result of the interaction of these two factors in the organization of information at the clause level. Referential structure may be realized in any particular language in different ways, usually word order and less commonly, case marking. The referential structure of a clause consists of at least one NP which is the referential peak, which is defined as the pragmatically most
salient NP due to either of the referentiality factors discussed above depending on the particular language, i.e. it may be given and/or definite in terms of the linguistic or extralinguistic context or it may be the highest ranking NP in the clause on the Referentiality Hierarchy. The referential structure of clauses in general represents the abstract organization of highly referential NPs within a clause. The referential structure of any particular clause represents the organization of the information bearing elements (especially the NPs) in such a way as to facilitate the communication of information between speaker and hearer. In English, the subject and direct object positions represent the abstract referential structure of the clause and the speaker's choice of which NPs occupy the subject and direct object positions reflect the speaker's intention to communicate with the hearer most effectively. In German, the nominative case performs a function analogous to that of subject position in English (Zubin 1976).

Thus far we have discussed only the organization of clause level grammar, which is but one level in the hierarchically organized system of grammar. These levels correspond to the levels of standard tagmemic theory: discourse, paragraph, sentence, clause, phrase, word and morpheme levels. In terms of our analysis, it is important to make the distinction between referential peak, which is a clause level phenomenon, and our notion of topic, which is a sentence level phenomenon. This distinction must be justified language specifically and will be motivated for each of the three languages to be discussed. However, there are at least two general criteria which may be suggested: (a) referential peaks bear selectional restrictions to their verb, whereas topics do not, and (b) there are clause level options such as the English passive which serve to rearrange the usual pattern of eligibility for the referential peak positions and which are signalled overtly by morphology or particles, whereas with topics there are merely variant word orders without any explicit clause level coding.

3. In this section we will present short typological sketches of the basic clause structure in each of the three languages to be investigated, starting with Tagalog.

3.1 Tagalog presents a unique type of organization of clause level grammar usually referred to in the literature as the focus system. Each NP in the Tagalog clause is marked for semantic role (ng for Actor, Patient and Instrument and sa for Locative, Goal, Source and Benefactive), except for one, which is the NP in focus. The NP in focus is the referential peak [RP] of the clause and is marked by ang. Only one RP per clause is allowed in Tagalog. Full NPs may generally occur in any order.

(1)a. B-um-ili ng isda sa bata ang lalake.  
AF-buy p fish S child RP man  
"The man bought some fish from the child."
    PF-buy A man S child RP fish
    "A/the man bought the fish from the child."

In (1a) the NP lalake 'man' filling the role of Actor is the RP of the clause and is marked by ang. The usual case marking for Actor ng is not present, but the -um- affixation in the verb indicates that the role of the NP marked by ang is the Actor. This type of clause is called an Actor-Focus(AF) construction. In (1b) the NP isda 'fish' is the RP marked by ang. The -in- infix in the verb indicates that it is functioning in the role of Patient. This is called a Patient-Focus (PF) or Goal-Focus construction. In (1c) the NP bata 'child' is the RP and is marked by ang. It fills the role of Source, and this is indicated in the verb by the suffix -an. This is called a Locative-Focus(LF) or Directional-Focus construction.

Thus, what is involved in all these examples is that a single NP is chosen as the RP of the clause, and its semantic role is indicated by affixation in the verb. This focus system is very rich and virtually any NP filling any semantic role may be the RP of the clause.

Note that all three examples have similar meanings except for differences in definiteness. NPs marked by ang are always definite, i.e. highly referential. Therefore, the factor determining eligibility of an NP to be the RP is the first of the two considerations for referentiality, context, linguistic and extralinguistic. We may see the referential structure in Tagalog as totally discourse controlled. Because the language has developed such an overt bifurcation of role functions and referential structures, we should expect a similar split in subject properties, which Schachter (1976, 1977) has in fact demonstrated with his split of subject properties into those controlled by the Actor and those controlled by what he terms the Topic, i.e. our RP, the NP marked by ang. The subject properties the RP has are referential properties, while those the Actor has are role properties.

There is a clear distinction in Tagalog between the clause level focus system and sentence level topicalization. If we compare the following with (1a-c):

(2) Sà tindahan b-um-ili ng isda ang lalake at b-in-asà
    L store AF-buy P fish RP man and PF-read

    niya ang diyaryo.
    he (A) RP newspaper

    'In the store the man bought some fish and read the newspaper.'

In (2) it is obvious that there is no special marking for the topic tindahan 'store' other than its sentence initial position, whereas the clause level RP is marked by the morpheme ang and more importantly by affixation in the verb. This affixation indicates that there is a semantic connection between the clause level RP and the verb which does not obtain between the sentence level topic and the verb.
Finally, Tagalog clauses are verb initial, whereas the sentence level topic is before the verb. The only exceptions to the verb initial typology of Tagalog are examples of sentence level topicalization such as above.

3.2. Navajo is an extremely complex and interesting language, and we will be concerned here only with a few features of clause level grammar. As in perhaps the majority of the world’s languages, the RP in the Navajo clause occurs clause initially. Unlike Tagalog, where discourse factors control which NP is the RP, in Navajo the **clausal initial NP** is determined to a large extent by the second of the two referentiality considerations, i.e. the inherent referentiality of the NP in terms of the Referentiality Hierarchy. In a clause containing an Actor and a Patient as full NPs, the NP in initial position must be the highest ranked NP on the Referentiality Hierarchy (Hale 1972):

(3a). \[ 'ashkii \ 'at’ééd \ yiyyiitsu’. \]
    boy(A) girl(P) saw

b. \[ 'at’ééd \ 'ashkii \ biiitsu’. \]
    girl(P) boy(A) saw
    'The boy saw the girl.'

(4a). \[ 'ashkii dził \ yiyyiitsu’. \]
    boy(A) mountain(P) saw

b. *dził \ 'ashkii \ biiitsu’.
    mountain(P) boy(A) saw
    'The boy saw the mountain.'

(5a). \[ *tó \ 'ashkii \ yiyyisxį’. \]
    water(A) boy(P) killed

b. \[ 'ashkii to \ biisxį. \]
    boy(P) water(A) killed
    'The water killed the boy.'

In (3) both NPs are of equal rank (human), and consequently, either NP may occur in initial position. The semantic role of each NP is coded in a prefix on the verb. When the first NP is the Actor and the second, the Patient, the verbal prefix is yi-, whereas when the Patient precedes the Actor, the verb takes the prefix bi-. In (4) the Actor outranks the Patient (human>inanimate), and therefore the only possible structure is (4a) with Actor-Patient-yi-verb. In (5), on the other hand, it is the Patient which ranks the highest (human>inanimate), and only (5b) with Patient-Actor-bi-verb word order is acceptable. We can summarize Hale’s (1972) findings as follows:

<table>
<thead>
<tr>
<th>Actor</th>
<th>Patient</th>
<th>Word Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>animate</td>
<td>animate</td>
<td>A-P-yi-verb</td>
</tr>
<tr>
<td>animate</td>
<td>inanimate</td>
<td>P-A-bi-verb</td>
</tr>
<tr>
<td>inanimate</td>
<td>animate</td>
<td>A-P-yi-verb</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P-A-bi-verb</td>
</tr>
</tbody>
</table>
It is obvious that these word order variants are clause level phenomena because they are controlled by clause internal considerations, namely, the inherent referentiality of the noun phrases themselves.

3.3. The two languages we have dealt with so far have had clearly definable referential structures in their system of clause level grammar. Lakhota, on the other hand, does not appear to have any such clause level referential structure, and we may therefore say that Lakhota is a role dominated language. By this we mean that the organization of clause level grammar is controlled by semantic roles and their interactions. Any referentially controlled options function at the sentence level and above (Van Valin 1977).

Semantic roles are coded in the verbal complex, and aside from instrumental NPs, full NPs are not explicitly marked for their semantical role. In a transitive sentence there is a set of pronominal prefixes which mark Actor, the third singular and plural of which is zero, and there is another set which marks Patients, the third singular of which is also zero. Because Lakhota is a stative-active language (Boas and Deloria 1939), some intransitive verbs take the Actor set of prefixes, while others use the Patient markers.

(6a. \( \emptyset \)-wa-kté.
P-I(A)-kill
'I killed it.'

b. wa-hí
I(A)-arrive
'I arrived.'

c. ma-\( \emptyset \)-kté.
me(P)-he(A)-kill
'He kills me.'

d. ma-híxpaye.
I(P)-fall down
'I fall down.'

e. mathó ki tháxca wā kté.
bear(A) the deer(P) a kill
'The bear killed a deer.'

In (6a), a transitive active clause, and (6b), an intransitive active clause, the Actor is marked by wa- 'I', and in (6a) the third person singular Patient is zero. In (6c), another transitive active clause, the third person singular Actor is zero, while the Patient is marked with ma- 'me', as it is in (6d), an example of an intransitive stative verb. Thus, first person singular is marked by wa- in these examples when it is an Actor and by ma- when it is a Patient; a similar dichotomy exists in second person pronouns. In the final example, (6e), both the Actor and the Patient have no overt marking of their semantic role, and the question of how these are determined arises. It appears that the first potential Actor NP in a Lakhota clause is the Actor. By 'potential Actor' is meant an NP which can fulfill the semantic role of Actor, i.e. its referent is
an animate sentient being capable of action. So the only way to specify the Actor is by word order. When there is only one potential Actor, then the NPs may occur in any order.

(7)a. ož̄až̄aglepi wã hokšila ki kablëče.
    window(P) a boy(A) the break
b. hokšila ki ož̄až̄aglepi wã kablëče.
    boy(A) the window(P) a break
    'The boy broke a window.'

Goals and Benefactives are also explicitly coded in the verbal complex, and here too there is no explicit marking on full NPs.

(8)a. wîchâša ki hokšila ki wówapi kipâzo.
    man(A) the boy(C) the book(P) show-to
    'The man showed the boy a book.'
b. wîchâša ki hokšila ki wówapi ophëkičithũ.
    man(A) the boy(B) the book(P) buy-for
    'The man bought the boy a book.'

In these examples the word order is crucial in determining semantic roles. The first human NP is the Actor and the second one the Goal or Benefactive; this Actor-Goal/Benefactive order is rigid, and only the position of the Patient wówapi "book" is variable. Instruments are marked on the verb by instrumental prefixes, and instrumental NPs are marked by the postposition ũ.

(9) mní khatá ũ žâžâ ki nablëče.
    water hot IN bottle(P) the break (internal force)
    'The hot water broke the glass.' or
    'He broke the glass with hot water.'

The instrumental prefix ka- on the verb in (7) means 'by striking', while the na- in (9) means 'by internal force.'

We have seen that word order functions in the Lakhota clause to distinguish the semantic role of NPs, since there is overt marking only on instrumental and locative NPs. We mentioned earlier that the two ways referential structure can be realized in a language are word order and case marking, but in Lakhota word order serves to differentiate semantic roles and there are no case markings. We may therefore conclude that there is no referential structure in the Lakhota clause. This does not mean, however, that discourse based referentiality considerations are ignored altogether; they function at the sentence level. This can be seen in the following example, in which discourse factors have overruled clause level semantic word order constraints when the role of the clause initial NP is identifiable from context.

(10)a. A: wîchâša ki taktókhũ he?
    man(P) the happen to question
    'What happened to the man?'
b. B: wičhaša ki thathaka ktep.
man(P) the buffalo(A) kill
'The man was killed by buffalo.'

In (10b) wičhaša 'man' is clearly a Patient and therefore may occur before thathaka 'buffalo', despite the fact that both NPs are potential Actors. (Out of context, (10b) would be ungrammatical, because the clause initial singular NP wičhaša would have to be the Actor, and the verb is marked for plural Actor, not plural (animate) Patient.) Thus, contextual factors at higher levels may override clause level constraints. There is no clause level coding in Lakhota which allows the NPs to be reversed and yet keeps semantic roles unchanged, as there is in English with its passive construction.

4. In spite of our earlier reservations about Keenan's (1976) attempt to provide a universal definition of subjects by means of his checklist of 30-odd properties, nevertheless some of these properties may be employed as diagnostics for investigating certain typological features of clause level grammar. In this section we will run through seven of these properties in each of the languages discussed. We will then attempt to demonstrate how basic typological features of each language as exemplified above controls the distribution of these properties.

4.1. Indispensability.

Keenan (1976:313) describes this property as follows: "A non-subject may often simply be eliminated from a sentence with the result still being a complete sentence. But this is usually not true of [basic]-
subjects. E.g. John hunts lions (for a living), John hunts (for a living)." Therefore, one of the major properties defining the subject of a clause in a language is its indispensability in Tagalog. Except for a few verbs such as meteorological verbs like umulan 'it rained,' all Tagalog clauses must contain an RP, i.e. an NP marked by ang. Therefore, regardless of its semantic role, the indispensable NP in a clause is the RP:

AP-buy RP man
'The man bought something.'

Pf-bite RP man
'Something bit the man.'

Navajo. The question of indispensability is more obscure in Navajo than in Tagalog. Going back to our discussion of yi- and bi- above, if the clause contains only one NP, it cannot be the RP of the clause, but rather must be the Patient with a verb with yi- and an Actor with bi-:

(12)a. 'at'eéd yiztaól
girl(P) kick
'He kicked the girl.'
b. 'at'éd biztač.
girl(A) kick
'The girl kicked him.'

Note in (12a) 'at'éd 'girl' is the Patient and obviously cannot be the RP of the clause, because yi- marks an Actor as RP. Likewise in (12b) 'at'éd is the Actor and cannot be the RP, because bi- marks a Patient as the RP. Thus, unlike in Tagalog, we cannot identify the RP of the clause as the indispensable NP.

Furthermore, sentences like (12a-b) present a formidable problem for any attempt to match up one's pretheoretical judgements of subjecthood with one's language specific notions of subject as identified by the methodology of the subject properties list. Compare (12a-b) with (13a-b).

(13a).
'ashkii 'at'éd yiztač.
boy(A) girl(P) kick
'The boy kicked the girl.'

b. 'at'éd 'ashkii biztač.
girl(P) boy(A) kick
'The boy kicked the girl.'

(13a-b) fits Keenan's definition of what constitutes a basic sentence in Navajo. Our pretheoretical judgements of what NP is the subject in (13a) is 'ashkii 'boy'. But the indispensable NP, as seen in (12a), is 'at'éd 'girl'. Our pretheoretical judgement is contradicted by the NP indispensability test. Furthermore, clauses like (13b) can be analyzed as analogous to an English passive construction in that the order of NPs is reversed in order to make an NP the RP of the clause. In such constructions, however, it is normally the referential peak NP which is indispensable, e.g. the clause-initial Patient is an English passive; and yet in Navajo it is the Actor, i.e. the NP which is not the RP, which is the indispensable NP. We may conclude from these facts that the application of the indispensable NP test to identify subjects in Navajo is very questionable.

Lakhota. The applicability of the indispensability test in Navajo has been shown to be problematic, and in Lakhota it collapses altogether. The problem with this test can be illustrated with the following example.

(14a).
Ø-Ø-kte
P A kill
'(he/she/it) killed (him/her/it)'

b. wícȟáša ki kte.
man (P) the kill
'He killed the man.'

As we pointed out earlier, third person singular Actor and Patient pronouns are zero in Lakhota, and therefore a 'naked' verb such as kte 'kill' can be a complete clause in and of itself. When we now ask about the indispensability of the NPs in such a clause, we encounter the following paradox: since there are no NPs in the sentence, we can conclude that there is in fact no indispensable NP in a basic Lakhota transitive clause; on the other hand, we could take the zero pronominal forms as somehow representing the indispensable NPs, in which case there
are two indispensable NPs in the clause. The paradox of indispensability in Lakhota is that the two possible interpretations of it yield either no indispensable NP or two such NPs, but never a unique NP as indispensable in a transitive clause. Furthermore, if we have a sentence such as (14b) with only one NP we find that it is interpreted as a Patient, rather than an Actor.

Given these data from the three languages, it is hard to see how the indispensability test can be defended as a test for subjecthood. While in Lakhota it fails to identify a unique NP, in Tagalog and Navajo it identifies two different types of NPs. In Tagalog it identifies the RP of the clause, while in Navajo it identifies the NP other than the RP, i.e. the NP of lower referentiality.

4.2 Coreferential Deletion Across Coordinate Conjunctions.

Keenan (1976:317) states "the NPs which can be coreferentially deleted across coordinate conjunctions include [basic] subjects."

(15) a. John went up to Fred and \( \emptyset \) insulted him.
    b. *John went up to Fred and he insulted \( \emptyset \).
    c. *John went up to Fred and he insulted \( \emptyset \).

To paraphrase Keenan, we may propose that subjects must fall into the class of NPs which can be coreferentially deleted across coordinate conjunctions.

Tagalog. In Tagalog we have already identified the RP of the clause marked by ang as the indispensable NP. The NP marked by ang may be coreferentially deleted across coordinate conjunctions, as may other NPs not marked by ang.

(16) a. Sa tindahan b-in-ili ng lalake ang diyaryo at
    L store PF-buy A man RP newspaper and
    b-in-asia niya ito.
    PF-read he (A) this
    'In the store the man bought the newspaper and read it.'

b. Sa tindahan b-in-ili ng lalake ang diyaryo at b-in-asia
    L store PF-buy A man RP newspaper and PF-read

In (16a) the RP is present in the second clause in the form of ito, the form of the deictic pronoun used for RPs (Schachter and Otanes 1972). (16b) is grammatical even though the RP of the clause has been deleted by coreferential deletion. Thus, in Tagalog the RP of the clause may be deleted coreferentially across coordinate conjunctions.

Navajo. Like in Tagalog, coreferential deletion across coordinate conjunctions in Navajo applies preferentially to RPs, but is not limited to them.

(17) a. 'ashkii ch'énádzid dôô 'at'êsëd vists'os.
    boy(A) woke up and girl(P) kissed
    'The boy woke up and kissed the girl.'

b. 'ashkii ch'énádzid dôô 'at'êsëd bists'os.
    boy(A) woke up and girl(A) kissed
    'The boy woke up and the girl kissed him.'
In both sentences the RP of the second clause has been deleted because it is coreferential with 'ashkii 'boy' in the first clause. (17a) cannot mean "The boy woke up and the girl kissed him"; in this sentence the deleted noun phrase must be the RP and yi- indicates it as an Actor. Similarly, (17b) cannot mean 'The boy woke up and he kissed the girl,' again because the deleted noun phrase must be the RP which is indicated as a Patient by the prefix bi-. Thus, in both sentences the deleted NP must be the RP peak.

However, while coreferential deletion applies preferentially to the RP, other NPs may also be deleted coreferentially in compound sentences. In such cases, the use of yi- and bi- function as a switch reference mechanism (Olson 1976; Jacobsen 1966), i.e. it monitors same versus different roles for the RPs in succeeding clauses:

(18) a. 'ashkii at'éd yi'íiitsá dóó yíztaž boy(A) girl(P) saw and kicked 'The boy saw the girl and he kicked her.'

b. 'at'éd 'ashkii biíitsá dóó yíztaž girl(P) boy(A) saw and kicked 'The boy saw the girl and she kicked him.'

c. 'at'éd 'ashkii biíitsá dóó bíztaž girl(P) boy(A) saw and kicked 'The boy saw the girl and he kicked her.'

d. 'ashkii 'at'éd yi'íiitsá dóó bíztaž boy(A) girl(P) saw and kicked 'The boy saw the girl and she kicked him.'

In all the sentences of (18), both 'ashkii 'boy' and 'at'éd 'girl' have been deleted coreferentially in the second clause. In (18a) yi- in the first clause marks the RP as Actor, and yi- in the second clause indicates the same; so we get 'he kicked her.' In (18b), bi- in the first clause indicates the Patient as the RP, i.e. 'at'éd 'girl', but the yi- in the second clause indicates that the RP, which must be 'at'éd 'girl' as in the previous clause, is the Actor; so we have 'she kicked him.' In (18d) bi- in the first clause indicates the RP is the Patient 'at'éd 'girl', and bi- in the second indicates that the RP is again a Patient; so we get 'he kicked her.' Finally, in (18d) yi- in the first clause indicates the Actor 'ashkii 'boy' as the RP, but bi- in the second clause indicates the coreferentially deleted RP 'ashkii must be a Patient; so we get 'she kicked him.'

Thus, what is important in all these examples is that the coreferentially deleted NP which is the RP is always the same as the RP of the first clause, and the change of yi- and bi- merely indicates the role of this NP. Consequently, in Navajo, like Tagalog, all coreferential deletions are keyed fundamentally on the RP of the clause which is the favoured NP for such options.

Lakhota. In introducing this subject property, Keenan (1976) presents several examples (see (15) above) in which a pronoun is deleted across a coordinate conjunction. In a language like English, the non-occurrence ('deletion') of an NP is not equivalent to the pronominalization
of that NP, and such a distinction underlies the applicability of this property for determining subjects in a language. We have already seen that in Lakhota third person pronouns are zero, and therefore non-occurrence ('deletion') is equivalent to pronominalization; consequently, this property is inapplicable in Lakhota as a test for subjecthood on an NP.

Since non-occurrence is pronominalization in Lakhota, it appears superficially as if 'deletion' across coordinate conjunctions is virtually unlimited.

(19) John khukhúše čépa ki čhí na ophéthu.
   A pig(P) fat the want and buy
   'John wanted the fat pig and (he) bought it.'

Here both the Actor and Patient occur as zero pronominal forms in the second clause. Oblique instrumental NPs have the same properties as Actor and Patient in this regard.

(20) a. wí ki ü ožaż¿glepi ki nabléche na čhága ská
   sun the IN window(P) the break and ice(P) melt
   'The sun broke the window and melted the ice.'
   b. wí ki ü čhága ki ská na hâhêpi wi ki ü akáxpe
   sun the IN ice(P) the melt and moon the INS cover
   'The sun melted the ice and the moon covered it (the sun).'

In (20a) wí 'sun' is an instrumental NP marked by the postposition ü which occurs in a zero pronominal form in the second clause as an Instrument, while in (20b) it is a pronominal Patient in the second clause. Thus, looking at coreferential deletion across coordinate conjunctions is not a possible diagnostic for subjects in Lakhota because of the identity of deletion and pronominalization. Like indispensability, it seems that this property is not workable as a universal test for subjects because in Tagalog and Navajo it identifies one type of NP, while in Lakhota it identifies no unique class of NPs.

4.3 Leftmost NP

Keenan (1976:319-320) states "[basic]-subjects are normally the leftmost occurring NPs in [basic]-sentences." This accounts for the rarity of VOS word order among the languages of the word.

Tagalog. This property is inapplicable to Tagalog because full NPs may occur in any order, although they all follow the verb.

Navajo. As can be seen clearly in the sentences with yi- and bi-
in 3.2 above, the RP is the leftmost NP in transitive clauses with two full NPs.

Lakhota. In our earlier general sketch of clause-level grammar in Lakhota, we saw that when there are two potential Actor NPs in a clause, the leftmost one is the Actor. (See (6e) above). Furthermore, where there is both an Actor and a human Goal or Beneficiary in a clause, the Actor must precede the Goal or Beneficiary (see (8) above). Thus in Lakhota the leftmost NP in a clause must be the Actor is there is a human or animate Patient, Goal or Beneficiary in the clause as well. This does not hold when the NPs filling
these other semantic roles are not potential Actors. (See (7) above.

4.4 Floating Quantifiers

Keenan (1976:320) discusses this property as follows: NPs which "launch" floating quantifiers (e.g. 'all the boys left' vs. 'the boys all left') include subjects. Keenan credits this observation to Postal and Perlmutter (1974), who claim that subjects preferentially launch floating quantifiers, then direct objects, and finally indirect objects. It is claimed that an implicational hierarchy exists such that indirect objects launch floating Qs, then direct objects and subjects must as well, and if direct objects can launch then, then so must subjects. Thus, if only one NP type in the clause may launch floating Qs in a particular language, then this would be identified as a subject in terms of this hierarchy. Furthermore, since subjects are the NPs preferentially launching floating Qs, the interpretation of one which is ambiguous between modifying the subject or direct object must in terms of this hierarchy be interpreted as modifying the subject.

Tagalog. In Tagalog a floating Q may only be interpreted as modifying the RP.

(21)a. B-um-ili-ng lahat ang mga lalake ng mga libro.
    AF-buy-lig all RP pl man P pl book
    'All the men bought some books.'

b. B-in-ili-ng lahat ng mga lalake ang mga libro
    PF-buy-lig all A pl man RP pl book
    'Some men bought all the books.'

In both clauses the floated Q lahat 'all' is interpreted as modifying the RP. (21a) is an Actor-focus construction, and so we have 'all the men bought some books', whereas (21b) is a Patient-focus clause meaning 'some men bought all the books.'

Lakhota. Quantifiers in Lakhota are normally taken to modify the NP immediately preceding them, as in (22).

(22)a. hokšila ki hená iyúha ožäžaglepi ki kabléblečhap
    boy(A) the pl all window(P) the shatter
    'All the boys shattered the windows'.

b. hokšila ki hená ožäžaglepi ki iyúha kabléblečhap
    boy(A) the pl window(P) the all shatter
    'The boys shattered all the windows.'

In (22) the quantifier iyúha 'all', modifies the immediately preceding NPs hokšila 'boy' in (22a) and ožäžaglepi 'window' in (22b).

Consider (23):

(23)a. hokšila ki hená ožäžaglepi ki xtálehã iyúha kabléblečhap.
    boy(A) the pl window(P) the yesterday all shatter
    'The boys shattered all the windows yesterday.'
b. hokšíla ki hená ečhákel iyúha ožážąglepi ki
boy(A) the pl intentionally all window(P) the
kabléblechapt.
shatter

'All the boys intentionally shattered the windows.'

The quantifier iyúha 'all' modifies the Patient ožážąglepi
Window(s)' in (23a) and the Actor hokšíla 'boy(s)' in (23b). The
quantifiers in these sentences can be said to be floating, since
they do not immediately follow an NP. Note in both cases the
quantifiers modify the NP closest to then on the left, not merely
the closest NP as is clearly demonstrated by (23b) in which
ožážąglepi is the NP closest to iyúha, but the quantifier modifies
hokšíla, the NP on its left. Note also that in (23a) the floated
quantifier modifies the Patient, not the Actor. The crucial test
for the Postal-Perlmutter hierarchy involves sentences in which
the interpretation of the quantifier is ambiguous; such an
example can be found in (24b).

(24a). wičhá-kte-p
them-kill-plural
'They kill them.'

b. iyúha wicháktep
all them-kill-plural
'They killed all of them.' ("They all killed them.")

In (24a) both the plural animate Actor and plural animate Patient
are explicitly coded in the verbal complex, and the only possible
reading is with the quantifier modifying the Patient and not the
Actor. An Actor reading is possible only when the Patient is singular
as in (25); in Lakhota, as in other languages, the NPs modified
by quantifiers must be plural.

(25) iyúha ktep
all kill-plural
'They all killed him.'

Thus in Lakhota the hierarchy for the interpretation of floating Qs
is Patient>Actor.

Looking at the data from Tagalog and Lakhota, we see that the
Postal-Perlmutter hierarchy fails to predict the facts in each
language. If we take the ang phrase in Tagalog as the subject,
then this hierarchy accounts for Tagalog but fails to deal with
the apparent Lakhota predilection for interpreting floating Qs
as modifying Patients (i.e. direct objects in transitive clauses
in relational grammar terms) over Actors. If, on the other hand,
we amend the hierarchy so that direct objects rank the highest
in order to account for Lakhota (and other languages such as Fijian
(Foley and Van Valin, in preparation)), then Tagalog becomes
problematic. In short, using facts relating to floating Qs as a
diagnostic test for subjecthood yields inconsistent and
contradictory results, and is therefore useless for defining
subjects universally.
4.5 Addressee of Imperatives

Keenan (1976:321) asserts that 'subjects' normally express the addressee phrase of imperatives.

**Tagalog.** The addressee is usually present in Tagalog imperatives, and it is the Actor regardless of whether it is the RP or not.

(26)a. B-um-ili ka ng isda sa bata
     AF-buy you-RP P fish S child 'Buy some fish from the child.'

     b. B-in-ili mo ang isda sa bata
     PF-buy you RP fish S child 'Buy the fish from the child.'

     c. Alis! 'Move'

     d. Um-alis ka! 'Move'

In (26a) the addressee of the imperative is both the Actor and the RP as indicated by the AF affix -um- in the verb. In (26b) and (26c) the addressee is the Actor but not the RP because (26b) is a Patient-focus clause. In (26c) the verb bears no focus affix, and so no Actor is required. (26d), on the other hand, has an affixed verb, and so the addressee must be present (here the RP). Thus, the addressee of imperatives in Tagalog is always an Actor.

**Navajo.** No data.

**Lakhota.** Imperatives are formed in Lakhota by the addition of the clitic -yo (or -wo depending on the final vowel of the preceding segment) by a male speaker or -ye (-we) by a female speaker to the end of the verbal complex. Since the addressee of an imperative in Lakhota must be a second person Actor with active verbs and a second person Patient with stative verbs, the Actor or Patient addressee is not marked on the verb.

(27)a. inážĩ-yo
     stand up-IMP
     'Stand up!'

     b. naxtáka-ye
     kick IMP
     'Kick it!'

     c. *nayaxtaka-ye
     you-kick IMP

     d. ištíma-yo
     sleep IMP (ištíma is a stative verb.)

     e. *ništíma-yo
     you-sleep IMP

While Actors are not marked in the imperatives of active verbs, Patients, Goals and Benefactives may be.

(28)a. mak'u-wo (Goal)
     me-give IMP
     'Give it to me!'
b. makte'-ye (Patient)
   me-kill IMP
   'Kill me!'

c. mi'iyuw'na-yo (Beneficiary)
   for me-bend IMP
   'Bend this for me!'

Thus, we may conclude that the addressee of an imperative of an active verb is always an Actor and that of a stative verb is always a Patient in Lakhota.

4.6 Reflexivization

Postal and Perlmutter (1974) claim that the same hierarchy which controls floating quantifiers is also applicable to the NPs controlling reflexivization. According to the hierarchy, subjects would preferentially be the controllers of reflexivization.

Tagalog. Actors, regardless of whether they are the RP or not, control reflexivization.

(29a). B-um-ili ng isda para sa kaniyang sarili ang lalake.
   AF-buy  P fish  B  him  self  RP  man
   'The man bought some fish for himself.'

b. I-b-in-ili ng lalake ng isda ang kaniyang sarili.
   perf-buy-LF  A  man  P  fish  RP  him  self
   'A/The man bought some fish for himself.'

c. *B-um-ili ng isda para lalake ang kaniyang sarili.
   AF-buy  P  fish  B  man  RP  him  self

d. *I-b-in-ili ng kaniyang sarili ng isda ang lalake.
   perf-buy-LF  A  him  self  P  fish  RP  man

In (29a) lalake 'man' is the RP and the Actor as indicated by the AF affix -um- in the verb and is controlling the reflexivization of the Benefactive. In (29a) lalake is the Actor which is controlling the reflexivization of the RP which is a Benefactive as signalled by the LF affix in the verb. The other two sentences, (29c) and (29d), are ungrammatical because in both cases a Benefactive controls the reflexivization of an Actor, where the Actor is the RP in (29) but not in (29d). Thus, in Tagalog reflexivization is always controlled by the Actor independent of whether it is the RP.

Navajo. With respect to reflexivization, Navajo patterns like Tagalog in that Actors control reflexivization regardless of whether they are the RP.

(30a). 'ashkii 'at'eed 'a-láąji' yiiitsá
        boy(A)  girl(P)  self-ahead  saw
        'The boy saw the girl in front of him(self)'

b. 'at'eed 'ashkii 'a-láąji' biïitsá
   girl(P)  boy(A)self  -ahead  saw
   'The boy saw the girl in front of him(self).'
In (30a) 'ashkii 'boy' is the referential peak and the Actor, as indicated by the prefix yi-, and controls the reflexivization of the oblique NP 'á-d-eé 'in front of him(slef)'. In (30b) 'at'éed 'girl' is the RP and is a Patient marked by bi-, but the Actor still controls the reflexivization of the oblique NP. Ashkii is again both the RP and Actor in (30c), and it controls the reflexivization of the oblique NP 'á-d-eé 'about himself'. In (30d) ashkii is not the RP but is the Actor, which controls reflexivization. Even though 'at'éed is the RP in (30d), it cannot control reflexivization, i.e. (30d) cannot mean 'the boy asked the girl about herself.' Lakhota. As in Tagalog and Navajo, reflexivization in Lakhota is controlled exclusively by Actors. Reflexive constructions are used to express the identity of Actor-Patient, Actor-Goal, or Actor-Beneficiary.

(31a). źúka ki ič'íslipe. 
dog(A) the self-lick  
'The dog licked himself.' (Actor-Patient)

b. heč'iye  
self-say that  
'He said that to himself.' (Actor-Goal)

c. John wówapi wá ophéič'ithū  
a book(P) a buy-for-oneself  
'John bought a book for himself.' (Actor-Beneficiary)

The reflexive pronoun ič'i- can occur only in the verbal complex, and therefore there can be no reflexivization into oblique NPs in Lakhota as is there in Navajo and English.

(32a). hokšíla ki wičhičala ki iwóič'iglake.  
boy(A) the girl(G) the talk-about-oneself  
'The boy told the girl about himself.'

b. Mary itómawapi wā wakípazo  
(G) they-took-a -picture-of-me (P) I-show-her/him  
'I showed Mary a picture of myself.'

Thus, in all three languages the Actor is the controller of reflexivization.

4.7 Relativization.
Keenan and Comrie (1972) discuss an implicational hierarchy of NP accessibility to relativization in various languages. They present the following hierarchy (Accessibility Hierarchy):

subject > direct object > indirect object > object of a preposition > genitive > object of comparative particle
If a language relativizes any category X on the hierarchy, then any category above X must be relativizable. Thus, if a language relativizes only one NP type, this must be that of subjects. Furthermore, any one relativization strategy utilized in a language must operate over a continuous segment of the hierarchy, although this may be only one category. Accordingly, a language may use one strategy for subjects and another strategy for other categories. German, for example, uses a participial relativization strategy for subjects but no other category. Relativization may therefore be used as a test for subjecthood in that if any NP type is relativizable in a language, then subjects may be.

Tagalog. Only the RP of the clause is relativizable in Tagalog. The referential peak, which is of course coreferential with the head noun, is deleted in the embedded clause.

(33a) isda-ng b-in-ili ng lalake sa bata
fish-lig PF-buy A man S child
'the fish that a/the man bought from the child'

b. *isda-ng b-um-ili ang lalake sa bata
fish-lig AF-buy RP man S child

c. *isda-ng b-in-ilh-an ng lalake ang bata
fish-lig perf-buy-LF A man RP child

d. bata-ng b-in-ilh-an ng lalake ng isda
child-lig perf-buy-LF A man P fish
'the child from whom a/the man bought some fish'

In (33a) we are relativizing on isda 'fish' which is a Patient in the embedded clause as indicated by the PF affix -in- in the verb; it is the RP of the lower clause and is therefore deleted. (33b) is ungram because we have relativized on the Patient isda 'fish', but the RP is lalake 'man' marked by ang; the AF affix in the verb indicates that the RP is the Actor. (33c) is ungrammatical as well because the relativized NP isda is not the RP of the clause. In (33d) we are relativizing on bata 'child' which is a Source in the embedded clause as signalled by the LF affix -an on the verb and which is deleted by virtue of its being the RP. Thus, we may conclude that in Tagalog only the RP of the clause may be relativized, regardless of its semantic role, and it is always deleted in the embedded clause.

Nava. According to Platero (1974), Navajo has two relativization strategies, the first involving the deletion of the head of the relative clause and the second the deletion of the RP of the lower clause which must be coreferential with the head. The latter strategy is of particular interest, because it applies exclusively to relativization on RPs, whereas the former appears to apply to any NP.

(34a) aíhosh-ígíi 'ashkkii aíháá'
sleep-REL boy(A) snore

b. 'ashkii aíhosh-ígíi aíháá'
boy(A) sleep-REL snore
'The boy who is sleeping is snoring.'
For the sake of this discussion, we will follow Platero (1974) in assuming that Navajo relative clauses are prenominal. (34a) illustrates the strategy in which the RP of the lower clause is deleted, while (34b) exhibits the head deletion strategy. Navajo gives preferential treatment to RPs in that it utilizes a relativization strategy which is applicable only to them.

(35a)  'at'éd biihtsá -(n)ee 'ashkii yáálti'.
girl (P) saw REL boy(A) speak
'The boy who saw the girl is speaking.'
(*'The boy who the girl saw is speaking.')

b. 'ashkii 'at'éd biihtsá -(n)ee yáálti'
boy(A) girl(P) saw REL speak
'The boy who saw the girl is speaking.' or
'The girl whom the boy saw is speaking.'

c. 'at'éd biihtsá -(n)ee 'ashkii yáálti'
girl(P) saw REL boy(A) speak
'The boy whom the girl saw is speaking.'
(*'The boy who saw the girl is speaking.' = (a))

d. 'ashkii 'at'éd biihtsá -(n)ee yáálti'
boy(P) girl(A) saw REL speak
'The boy whom the girl saw is speaking.' or
'The girl who saw the boy is speaking.'

In the first example, the RP of the embedded clause 'ashkii 'boy' has been deleted, and it must be an Actor because the verb is marked with yá- . Thus the only possible interpretation is 'the boy who saw the girl is speaking', where 'boy' is the Actor in the relative clause. The second interpretation is unacceptable because in it 'boy' is a Patient and this is incompatible with the yá- prefix. Example (35b) is ambiguous because the head has been deleted and this strategy is not limited to RPs. The yá- prefix indicates that the RP 'ashkii 'boy' is an Actor, but the deleted head may be either 'ashkii or 'at'éd 'girl'; consequently the sentence is ambiguous between readings with 'ashkii or 'at'éd as the head. The ambiguity accrues from the fact that this strategy is not limited to the RP of the lower clause, allowing either NP to be interpreted as the head. In (35c) the RP is again 'ashkii, and in this case it must be the Patient as indicated by the bi- prefix on the verb. Thus the meaning 'the boy whom the girl is speaking' is the only possible one because the RP deletion strategy has been followed here, which means that only 'ashkii can be interpreted as the head noun. In the final example we again have the head deletion strategy resulting in ambiguity. Because the verb is marked by bi-, 'ashkii must be the Patient, regardless of which NP is the head. We may say, then, that Navajo gives special attention to relativization on RPs in that it possesses a relativization strategy restricted to them.

Lakhotā. Lakhotā relativization differs from that of both Tagalog and Navajo in that it is not restricted to one particular NP type as in Tagalog but is virtually unconstrained and in that unlike Navajo is has only one relativization strategy which applies to all
NP types. Any NP may be relativized upon in Lakhota regardless of its semantic role or syntactic function.

(36)a. wičhaša wā šūka naxtače ki he thaló yúte. (Actor) man (A) a dog (P) kick the meat (P) eat 'The man who kicked the dog eats meat.'

b. wičhaša wā xtaleha wablāke ki he wówapi man a yesterday I-see the book (P) ophēthū. (Patient) buy 'The man I saw yesterday bought a book.'

c. wičhaša wā wiyā ki wówapi k'ū ki he thaló yúte. (Goal) man a woman the book give the meat (P) eat 'The man to whom the woman gave the book eats meat.'

d. čhā wā ū šūka ki awāpane k'ū he wičhaša stick a INST dog the I-hit the man (A) ki yuwēge. (Instrument) the break 'The man broke the stick with which I hit the dog.'

e. wičhaša wā thikí xugnáge k'ū he čhiyéwayne. (Genitive) man a whose-house burn-down the my-brother 'The man whose house burned down is my brother.'

f. wiyā ki wičhaša wā isâmya háske k'ū he woman the man a than tall the čhiyéwayne. (Obj. of Comp. Part.) my-brother 'The man who the woman is taller than is my brother.'

We will follow Van Valin (1977) in analyzing Lakhota relative clauses as being postnominal. The basic structure of the relative clauses in these sentences is a sentence in which the NP coreferential with the head has been deleted and which has been nominalized by the articles ki 'the' or k'ū 'the aforementioned'. The crucial difference between relativization in Navajo and Lakhota is that Navajo has one strategy which applies only to RPs and another one which applies to any NP regardless of whether it is an RP, whereas Lakhota has only one strategy which is equally applicable to all NPs.

To make sense of these data from the three languages, it is important to realize that relativization is fundamentally a referential phenomenon, i.e., a sentence is appended to an NP in order to further delimit its reference. Regardless of whether we assume the traditional analysis of relativization or Schachter's (1973) reanalysis, it is straightforward that the NP in the embedded clause which is being relativized upon is highly referential. Consequently, we should expect to find languages which restrict relativization entirely to the RP of the embedded clause, e.g. Tagalog, and languages which have developed relativization strategies which single out the RP of the embedded clause for special attention, e.g. Navajo. However, a language which has no referential structure
(i.e. heavily role dominated language) should not have any such restrictions on relativization. Thus, this account of relativization makes sense of the Lakhota facts because we have claimed that it is a role dominated language on other grounds, and the facts relating to relativization give further support to our claim. Lakhota does not distinguish among NPs in relativization strategies because there is no RP in the Lakhota clause. All roles may be relativized upon, and the same strategy is used in all cases. For a more complete discussion of this theory and its implications for Keenan and Comrie's (1972) Accessibility Hierarchy, see Foley (in preparation).

5. We can summarize the preceding discussion in the following chart:

<table>
<thead>
<tr>
<th>SUBJECT PROPERTY</th>
<th>TAGALOG</th>
<th>NAVAJO</th>
<th>LAKHOTA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indispensability</td>
<td>RP</td>
<td>non-RP</td>
<td>All or none(Patient?)</td>
</tr>
<tr>
<td>Coreferential Deletion</td>
<td>RP</td>
<td>RP</td>
<td>Unrestricted</td>
</tr>
<tr>
<td>Relativization</td>
<td>RP</td>
<td>RP(restricted)Unrestricted</td>
<td></td>
</tr>
<tr>
<td>Imperative Addressee</td>
<td>Actor</td>
<td>No data</td>
<td>Active vb:Actor</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Stative vb: Patient</td>
</tr>
<tr>
<td>Reflexivization</td>
<td>Actor</td>
<td>Actor</td>
<td>Actor</td>
</tr>
<tr>
<td>Leftmost NP</td>
<td>Inapplicable</td>
<td>RP</td>
<td>Actor&gt; Patient</td>
</tr>
<tr>
<td>Floating Quantifier</td>
<td>RP</td>
<td>No data</td>
<td>Patient&gt;Actor</td>
</tr>
</tbody>
</table>

We will first discuss the relationships holding among various properties and then analyze their distribution in each language in light of its typological features. The two most clearly role related properties are reflexivization and imperative addressee, both of which are connected to role function and not referential structure in all three languages. In the languages we have examined, the addressee of an imperative is keyed to the semantic role of Actor (with the exception of stative verbs in Lakhota), although some languages may add the further restriction that it be the referential peak as well. Reflexivization is associated both with semantic roles such as Actor and Patient as well as with the linear order of constituents, (e.g. Samoan (Chapin 1970)). Thus, both of these 'subject' properties are associated in universal grammar with the role functions of the clause.

The rest of the properties are connected in universal grammar with the referential structure of the clause. These properties do not have any inherent interrelationships in universal grammar but rather are organized in terms of language specific typological facts. However, relativization and coreferential deletion across coordinate conjunctions appear to be related, as can be seen from their parallel behavior in Tagalog, Navajo, Lakhota, and most strikingly Dyirbal (Dixon 1972). The relationship between leftmost NP and floating quantifiers is a complex one mediated by typological considerations of word order and referential structure. The leftmost NP is the main RP in the vast majority of the world's languages, and therefore when the quantifier is floated towards the verb in verb-initial or verb-medial languages, it is not surprising that we find that the quantifier is interpreted as modifying (having been launched by) the nearest NP which is the leftmost NP and which is most likely to be
the RP. However, in verb-final languages this situation does not obtain, since when the quantifier is floated towards the verb, the nearest NP is not the leftmost one.

We now turn our attention to a discussion of these inter-relationships within a language-specific typological perspective. Tagalog, as we have seen, has a strict separation between the role and referential features of clause level grammar. Consequently, the role affiliated properties of imperative addressee and reflexivization are restricted to the semantic role of Actor regardless of its referential features, whereas the referential properties of indispensability, coreferential deletion, floating quantifiers, and relativization are restricted to the RP irrespective of its role function. Because Tagalog is a verb-initial language with free word order in which the RP is marked by a preposition (ang), there can be no direct association between leftmost NP (which is irrelevant at the clause level in Tagalog) and floating quantifiers, and the latter are always interpreted as modifying the NP marked by ang, i.e. the RP.

Although our data are not as complete for Navajo, it appears to show the same general dichotomy between role and reference properties. Reflexivization again seems to be associated with the semantic role of Actor and not the RP, whereas coreferential deletion, leftmost NP and relativization are consistently keyed to the referential structure. The leftmost NP is always the referential peak of the clause, and coreferential deletion and relativization give preferential treatment to the RP.

If we compare the properties listed in the Lakhota column with those in the Tagalog and Navajo columns, one is struck by their unusual behaviour. There are no properties related to anything like the referential structures we discussed in Tagalog and Navajo. There are nevertheless connections between the properties, albeit on a different basis. Reflexivization is, as in the other languages, controlled by the semantic role of Actor, and the addressee of an imperative of an active verb is likewise keyed to this semantic role, although the imperatives of stative verbs take Patients as addressees. The other five properties are referentially based, but we have shown Lakhota to be a language without a referential structure to which such properties can be related. Indispensability is a problematic property in Lakhota, since, as we showed earlier, it fails to pick up a unique NP on one interpretation, and identifies Patients rather than Actors on the other. Like indispensability, coreferential deletion is vitiated as a diagnostic test for subjeckthood because of the equivalence of deletion and pronominalization. Furthermore, the lack of any referential structure precludes the possibility of selectively attending to particular NPs both for coreferential deletion and relativization, which applies freely to all NPs and employs the same strategy in all cases. The final two properties, leftmost NP and floating quantifiers, are characterized by inverse hierarchies which express the relationship between quantifier float interpretation and verb-final word order noted above: in Lakhota the unmarked word order is Actor-Patient, and so when a quantifier is floated towards the verb, the NP closest to it will usually be the Patient, thus resulting in the preferential interpretation of the quantifier as
modifying the Patient rather than the Actor. Therefore, we see that the behaviour of these properties in Lakota is intelligible only in terms of our analysis of it as a role dominated language with no referential structure.

6. The notion of subject has played three major roles in linguistics. It has been used as an explanatory concept, a descriptive device, and a theoretical construct. An example of the use of 'subject' as an explanatory concept is in Keenan and Comrie's (1972) Accessibility Hierarchy in which 'subject' functions as the highest category in order to explain relativization phenomena in languages like Malagasy and Tagalog. As a descriptive device, the notion has functioned as a label in the descriptions of many languages, such as the analysis of the English clause into 'subject and predicate'. Finally, 'subject' has been identified as a theoretical construct in theories like relational grammar in which it is assumed as primitive.

Keenan (1976:305) points out the need for an explicit definition of the notion of subject as a theoretical construct in order to ground its validity as an explanatory concept in many of the generalizations of universal grammar.

Clearly generalizations [such as the Accessibility Hierarchy] determine constraints on the form, and substance, of possible human languages. But to verify them and determine their universality it is necessary to be able to identify subjects, direct objects, etc., in a principled way across [languages]. If we use different criteria to identify subjects in different [languages] then "subject" is simply not a universal category and apparently universal generalizations stated in terms of that notion are not generalizations at all. (1976:305 [emphasis added])

We have taken the criteria which Keenan proposes and have shown that they fail to identify a unique NP as the 'subject' in each language, e.g., they identify both the RP and the Actor as potential subjects in Tagalog. Furthermore, the same property does not consistently identify the same NP in different languages, e.g., indispensability picks out the RP as subject in Tagalog but the non-RP as subject in yi- and bi-clauses in Navajo, while in Lakhota it fails to identify any unique NP at all as subject; and this undermines Keenan's attempt to define 'subject' as a "cluster concept", since the same properties cluster around different NPs in different languages. Therefore, the attempt to elevate the notion of subject to a theoretical construct is doomed, because no explicit universal definition can be given, i.e. no set of criteria can be given which will consistently identify the same NPs as subjects not only cross-linguistically but also within some languages. Consequently, it loses its value as an explanatory construct, thus vitiating such proposed universal generalizations as the Accessibility Hierarchy in their present form (see Foley in preparation).

At this point one might object that our arguments apply only to attempts to define subjects at one level, namely, surface subjects,
and that the facts we have cited can be accommodated within a theory (e.g. relational grammar) which posits subjects at more than one level, i.e. deep and surface subjects. In particular, Perlmutter and Postal (1977) claim that Actors are "initial subjects" in the underlying form of a sentence (initial stratum in a relational network) and what we have called the RP of a clause would be the "superficial (or cyclic) subject". We will accept these equivalences for the sake of the argument, keeping in mind that they are not in fact equivalent; RPs are defined in purely pragmatic terms (see section 2), while superficial subjects are claimed to have both role (e.g. control reflexivization) and referential (e.g. relativizable) properties. We may accordingly recast the table presented at the beginning of section 5 in terms of initial and superficial subjects.

<table>
<thead>
<tr>
<th>INDISPENSABILITY</th>
<th>TAGALOG</th>
<th>NAVAJO</th>
<th>LAKHOTA</th>
</tr>
</thead>
<tbody>
<tr>
<td>COREF. DELETION</td>
<td>superficial</td>
<td>initial?</td>
<td>superficial?</td>
</tr>
<tr>
<td>RELATIVIZATION</td>
<td>superficial</td>
<td>superficial</td>
<td>any NP</td>
</tr>
<tr>
<td>IMPERATIVE ADD</td>
<td>initial</td>
<td>initial</td>
<td>initial</td>
</tr>
<tr>
<td>REFLEXIVATION</td>
<td>initial</td>
<td>initial</td>
<td>initial</td>
</tr>
<tr>
<td>LEFTMOST NP</td>
<td>--</td>
<td>superficial</td>
<td>initial</td>
</tr>
<tr>
<td>FLOATING Q</td>
<td>superficial</td>
<td>--</td>
<td>superficial?</td>
</tr>
</tbody>
</table>

The question marks next to several of the entries must be clarified before we go any farther. There is a question mark next to the indispensable NP in Navajo: we saw earlier that in vi- and bi-clauses the indispensable NP is the non-referential peak NP, and while this is the initial subject (Actor) in a bi-clause, it is neither the initial nor superficial subject in a vi-clause. Both constructions are problematic for relational grammar. As we saw in (12a), the indispensable NP in a vi-clause is the Patient (direct object), and there is no way to make that NP a subject of any kind, since the rule that would do so creates bi-clauses as in (3b) and (12b). Such constructions meet the definition of a passive clause given in Perlmutter and Postal (1977), i.e. the NP in the "corresponding" active sentence which is the initial direct object (Patient), is the superficial subject (RP) in the derived "passive" clause. However, such sentences contradict the three major universal claims Perlmutter and Postal make about passivization: (1) they claim that the direct object of the active is the superficial subject of the corresponding passive, but in (12b) the former direct object is not the superficial subject, as it has been deleted; (2) they claim further that the subject of an active clause is neither the superficial subject nor direct object of the corresponding passive, and yet in Navajo the Actor (demoted initial subject) in a bi-clause still controls reflexivization (see Shachter (1977:30ff) for discussion of the similar situation in Tagalog) and is the indispensable NP, thereby exhibiting important subject-like properties and contradicting the Chômeur Condition (n.e. the Relational Annihilation Law); and (3) on the basis of the first two generalizations they claim that passive clauses are superficially intransitive, and (12b) is extremely problematic for such a claim.
since it is not only semantically transitive but the superficial subject has been deleted, an impossibility in true intransitive clauses. Thus, the facts regarding NP indispensability in vi- and bi-clauses in Navajo cannot easily be brought into line with the generalizations which relational grammar seek to make about initial and superficial subjects.

There are also question marks next to the 'superficial' subjects in the indispensability and floating quantifier rows in the Lakhota column, because these are properties associated with Patients, not Actors. Consequently, in order to express these Lakhota facts in terms of the subject categories we have been using, a relational grammarian would have to claim there is some sort of advancement rule in Lakhota which turns underlying Patients (direct objects) into superficial subjects, i.e. a passive rule. The problem with this solution is that there is absolutely no morphological, syntactic or semantic evidence whatsoever for the existence of such a rule in Lakhota (see Van Valin 1977), and moreover, to postulate such a rule solely on the basis of the proposed universals one is trying to make the data fit, is not only unmotivated and circular but also adds unnecessary complexity to the grammar of Lakhota. Nevertheless, to account for (24b) with such a rule and thereby to preserve the Postal-Perlmutter (1974) hierarchy for floating quantifiers, one would have to claim: (1) in the underlying form there are two NPs, a subject and a direct object, which are so marked on the verb (cf. (24a)), plus the quantifier iyuha 'all'; (2) the direct object to subject rule applies, making the direct object a subject but not changing the subject and object agreement marking on the morphologically unchanged verb (recall that grammatical function is determined by NP order and role coding in the verb (see 3.3)); (3) the quantifier is now interpreted as modifying (having been launched by) the derived subject which is still marked as an object on the verb; and (4) both NPs are deleted. Here we encounter significant if not insurmountable difficulties in expressing the Lakhota facts in terms of initial or superficial subjects, and it appears that not even two kinds of subjects are adequate for expressing the desired generalizations in this case.

Looking back at the table we drew up in terms of initial and superficial subjects, we must now ask, what kinds of universal generalizations about subjects can we make? This is a crucial question, because Postal and Perlmutter (1974) do not attempt to define 'terms', i.e. subjects, direct objects and indirect objects, but rather explicate them in terms of the grammatical processes in which they are involved. Thus, to say that an NP type in a language is a subject means in their terms that it behaves in accordance with the generalizations established by the theory about the behavior of subjects, e.g., if only one kind of NP in a language can trigger reflexivation and launch floating quantifiers, then that NP type is a subject at some level. It should be noted that talking about subjects on more than one level already weakens the notion significantly in that generalizations about subjects must be stated in terms of one subset of subjects or the other and not about subjects in general. There are only two straightforward and unproblematic generalizations about
subjects that can be made: (1) initial subjects (Actors) are the addressee of imperatives (of active verbs) and (2) superficial subjects (RPs) are always subject to relativization, and in some cases they may be the only NPs which may be relativized. There appears to be a generalization that initial subjects control reflexivization, but this contradicts the claim made by Postal and Perlmutter (1974) that superficial subjects control reflexivization. Consequently, since there are cases in other languages where superficial subjects trigger reflexivization, any generalization about subjects and reflexivization must be stated in terms of any kind of subject (initial or superficial) being able to trigger it, and this is much weaker than their earlier generalization.

The situation gets worse with respect to the other four properties. We have already seen the immense difficulties attending the interpretation of the indispensable NP in Navajo and Lakhota in terms of subject categories, and even if one managed to state the facts in terms of deep or surface subjects, the generalization one could make would be extremely weak: the indispensable NP is some kind of subject. The generalization is so weak, in fact, that it borders on circularity: in order to interpret the indispensable NP in Navajo and Lakhota as a subject, one must postulate rules whose only motivation is to make these NPs subjects in order to be able to fit the generalization that the indispensable NP is some kind of subject. While there appears to be some generalization about coreferential deletion, namely that superficial subjects are among the NPs that may trigger it, the heart of the generalization, which is that only terms may trigger it (Postal and Perlmutter 1974), is contradicted by examples (20a) and (20b) from Lakhota in which an oblique NP is triggering coreferential deletion across conjunctions. This not only vitiates the generalization that only terms may trigger it, but also the generalization about subjects, since if any NP can trigger coreferential deletion, then saying that subjects are among the NPs which can is vacuous. With respect to leftmost NP, one must here again say that "some kind of subject" is the leftmost NP in a sentence. And finally, in order to make a coherent generalization about launching floating quantifiers, one must again resort to the postulation of otherwise unmotivated rules in Lakhota in order to interpret the Lakhota facts in terms of some notion of subject.

The answer to our question about the kinds of generalizations we can make about subjects is unambiguous: in the cases where generalizations are stateable, they are extremely weak, and in several cases even the weak generalizations are not stateable without the circular postulation of otherwise unmotivated rules in one or more of the languages. In the two cases where apparently valid generalizations are possible, each refers to a different kind of subject. In no case do we have several significant and substantial generalizations referring to one type of subject, much less to both types. Furthermore, in all of these cases (e.g. relativization, reflexivization, and imperative addressee), the relevant generalizations can be stated without any reference to the notion of subject at all, as we showed above (see section 4 and table at the beginning of section 5.)
one nevertheless wishes to maintain that subject is a viable universal category, then the burden of proof is on him to show that despite the problems we have exhibited, 'subject' can be defined universally in some way and that significant generalizations can be stated in terms of subjects so defined. We have shown that all previous major attempts to define subject universally have failed and furthermore that only extremely weak and in some cases vacuous generalizations can be stated about subjects on any level, thereby vitiating the attempt to define them solely in terms of the grammatical processes in which they are involved. Consequently, our conclusion that subject is not a valid theoretical construct (universal) in linguistic theory stands.

In this section we have discussed the three main uses of the notion of 'subject' in linguistics, i.e. as an explanatory concept, a theoretical construct, and a descriptive device. Since the value of the notion as an explanatory concept depends crucially on its validity as a theoretical construct, our arguments against its universality undermine its explanatory role as well. Thus, we are left with 'subject' as a potentially useful tool for the description of some languages, e.g. English and French. It must be borne in mind that when using 'subject' as a descriptive label, it is just that and nothing more; it may label a different entity in the grammar of any language to which it is applied. Any attempt to claim that the NPs designated by this term in two different languages are equivalent raises the notion to the status of a theoretical construct which we have shown to be invalid.

7. We began this paper by noting that the notion of subject has figured prominently in grammatical discussions since the time of the Greeks. This notion has come down to us by two paths: through the grammatical tradition which has taken the subject-predicate dichotomy as basic, and through traditional logic, which also takes this distinction as its starting point. This dichotomy is the mold into which most grammatical theorising and description in the Western tradition has been forced. By rejecting the notion of subject as a construct in linguistic theory, we have taken the first step in the "task of liberating grammar from logic." (Heidegger 1927:209)

FOOTNOTES

1. We would like to thank our native speaker consultants for their assistance and their patience: Ms. Teresita Zaragoza and Mrs. Shirley Yengoyan (Tagalog), Mrs. Eva Brown (Lakhota), and Mr. Dave White (Navajo). We would also like to thank Lawrence Reid, George Lakoff, and Anthony Woodbury for helpful discussions of some of the points covered in the paper; they are not responsible for anything said herein. The following abbreviations are used in the paper:

<table>
<thead>
<tr>
<th>A</th>
<th>Actor</th>
<th>IN</th>
<th>Instrument</th>
<th>RP</th>
<th>Referential Peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>Patient</td>
<td>AF</td>
<td>Actor Focus</td>
<td>lgr</td>
<td>Ligature</td>
</tr>
<tr>
<td>S</td>
<td>Source</td>
<td>PF</td>
<td>Patient Focus</td>
<td>pl</td>
<td>Plural</td>
</tr>
<tr>
<td>B</td>
<td>Beneficiary</td>
<td>LF</td>
<td>Locative Focus</td>
<td>Q</td>
<td>Quantifier</td>
</tr>
</tbody>
</table>
REFERENCES


______ and R. Van Valin. in preparation. Language and the Sciences of Man.


