In section 1 we present the Accessibility Hierarchy, in terms of which we state three universal constraints on Relative Clause Formation. In addition, we present the data in support of these constraints and discuss certain partial counterexamples. In section 2 we propose a partial explanation for the hierarchy constraints and present further data from Relative Clause Formation supporting these explanations. Finally, in section 3 we refer briefly to other work suggesting that the distribution of advancement processes such as Passive can be described in terms of the Accessibility Hierarchy; these facts show that the proposed explanation for the Hierarchy needs to be generalized.

1. The Accessibility Hierarchy

1.1. Two Methodological Preliminaries

We are attempting to determine the universal properties of relative clauses (RCs) by comparing their syntactic form in a large number of languages. To do this it is necessary to have a largely syntax-free way of identifying RCs in an arbitrary language.

Our solution to this problem is to use an essentially semantically based definition of RC. We consider any syntactic object to be an RC if it specifies a set of objects (perhaps a one-member set) in two steps: a larger set is specified, called the domain of relativization, and then restricted to some subset of which a certain sentence, the

* This article is a developed version of Keenan and Comrie (1972). We attempt in this version to account for several objections to the earlier formulation that have since been brought to our attention.

We would like to acknowledge the following sources for significant help with this article:

(i) One Linguistic Inquiry reviewer for constructive criticisms of an earlier draft;

(ii) Support from an NSF postdoctoral research grant and two Wenner-Gren grants, 2384 and 2944, for work on the Malayo-Polynesian languages reported herein; for support of the psycholinguistic work on the Accessibility Hierarchy we are indebted to a grant from the Nuffield Foundation;

(iii) The willing and substantive help from the following linguists concerning languages they have worked with: G. Brettschneider (Basque), P. Brown (Tzeltal), M. Butar-Butar (Toba Batak), S. Chung (Indonesian, Tongan), O. Dahl and J. Allwood (Swedish), K. Ebert (North Frisian, Kera), J. Hawkins (German, Slovenian), A. Janhunen (Finnish), M. Perera (Catalan), H. van Riemsdijk (Zurich German), J. de Rooij (Dutch), A. Salmond (Maori), P. Sgall (Czech), R. Tanaka (Japanese), B. Vattuone (Genoese), N. Vincent (Italian);

(iv) The willing and substantive help of our informants.
restricting sentence, is true.¹ The domain of relativization is expressed in surface structure by the head NP, and the restricting sentence by the restricting clause, which may look more or less like a surface sentence depending on the language.

For example, in the relative clause the girl (that) John likes the domain of relativization is the set of girls and the head NP is girl. The restricting sentence is John likes her and the restricting clause is (that) John likes. Clearly, for an object to be correctly referred to by the girl that John likes, the object must be in the domain of relativization and the restricting sentence must be true of it. We shall refer to the NP in the restricting sentence that is coreferential with the head NP as the NP relativized on (NP_rel); in our example, this is her, i.e. the direct object of John likes her.

Note that we only consider definite restrictive RCs in this study. The role of the determiner the is held constant and ignored, and the term RC is used to apply to the collocation of the head NP and the restricting clause.

Note further that our semantically based notion of RC justifies considering as RCs certain constructions that would perhaps not have been so considered in traditional grammar. Thus, in German, alongside the traditional RC in (1) we also count the participial construction in (2):

1) der Mann, der in seinem Büro arbeitet
   the man who in his study works
   ‘the man who is working in his study’

2) der in seinem Büro arbeitende Mann
   the in his study working man
   ‘the man who is working in his study’

As the German data above illustrate, not only do different languages vary with respect to the way RCs are formed, but also within a given language there is often more than one distinct type of RC. We shall refer to distinct ways of forming RCs as different relative clause forming strategies. Different strategies differ with regard to which NP positions they can relativize. Thus, the participial strategy in (2) above can only relativize subjects (that is, the head NP can only be understood to function as the subject of the main verb of the restricting clause), whereas the strategy in (1) above functions to relativize almost any major NP position in simplex sentences. Consequently, generalizations concerning the relativizability of different NPs must be made dependent on the strategies used. It will be critical therefore to provide some principled basis for deciding when two different RCs have been formed with different strategies.

There are many ways RCs differ at the surface, and hence many possible criteria for determining when two strategies are different. We have chosen two criteria that seem to us most directly related to our perception of how we understand the meaning of the RC—that is, of how we understand what properties an object must have to be

¹ A more formal statement of this semantic notion in which the logical structure of an RC is represented as a pair, a common noun phrase and an open formula, can be found in Keenan (1972a).
correctly referred to by the RC. The first concerns the way the head NP and the restricting clause are distinguished at the surface, and the second concerns how the position relativized is indicated.

In the first case we consider two RCs to be formed by different strategies if the relative position of the head NP and the restricting clause differs. There are three possibilities: the head occurs to the left of the restricting clause, as in (1) above (postnominal RC strategy); the head occurs to the right, as in (2) (prenominal RC strategy); or the head occurs within the restricting clause (internal RC strategy), as in (3) and (4), from Bambara (Bird (1966)) and Digueño (Gorbet (1972)), respectively:

(3) a. ne ye so ye.
   I Past horse see
   'I saw a horse.'

b. ne ye so min ye
   I Past horse which see
   'the horse that I saw'

c. tye ye ne ye so min ye san.
   man Past I Past horse which see buy
   'The man bought the horse that I saw.'

   yesterday house DO I-saw
   'I saw the house yesterday.'

b. ?əwa:+ pu + Lyw ciyawx.
   house Def in I-will-sing
   'I will sing in the house.'

   yesterday house DO I-saw Def in I-will-sing
   'I will sing in the house that I saw yesterday.'

In the second case we consider two RCs to be formed from different strategies if one presents a nominal element in the restricting clause that unequivocally expresses which NP position is being relativized, and thus we know exactly what the restricting clause is saying about the head NP (that is, we can recover the restricting sentence from surface) (+case RC strategy). For example, the English strategy that forms the girl who John likes is not case-coding since who, the only relevant particle in the restricting clause, can be used as well if the role of the head NP in the restricting clause is different, e.g. the girl who likes John (−case RC strategy). On the other hand, in comparable sentences in Russian, (5a) and (5b), the form of the relative pronoun does unequivocally tell us the role of the head NP, so that strategy in Russian is case-coding:

(5) a. devuška, kotoruju Džon ljubit
   girl who ( accusative) John likes
   'the girl who John likes'
b. devuška, kotoraja ljubit Džona
   girl who (nominative) likes John
   'the girl who likes John'

Note, however, that RCs in English like the chest in which John put the money are considered case-coding, since the preposition in, which indicates the role of the head NP, is present in the restricting clause.

In addition to the use of relative pronouns, case can be coded in another way in the languages covered in our study. Namely, a personal pronoun can be present in the NP position relativized, as in Hebrew:

(6) ha- isha she- Yon natan la et ha- sefer
    the woman that John gave to-her DO the book
    'the woman that John gave the book to'

1.2. The Accessibility Hierarchy and the Hierarchy Constraints

1.2.1. Statement of the Hierarchy and the Constraints. On the basis of data from about fifty languages, we argue that languages vary with respect to which NP positions can be relativized, and that the variation is not random. Rather, the relativizability of certain positions is dependent on that of others, and these dependencies are, we claim, universal. The Accessibility Hierarchy (AH) below expresses the relative accessibility to relativization of NP positions in simplex main clauses.

   Accessibility Hierarchy (AH)
   SU > DO > IO > OBL > GEN > OCOMP

Here, ‘>’ means ‘is more accessible than’; SU stands for ‘subject’, DO for ‘direct object’, IO for ‘indirect object’, OBL for ‘major oblique case NP’ (we intend here NPs that express arguments of the main predicate, as the chest in John put the money in the chest rather than ones having a more adverbial function like Chicago in John lives in Chicago or that day in John left on that day), GEN stands for ‘genitive’ (or ‘possessor’) NP, e.g. the man in John took the man’s hat, and OCOMP stands for ‘object of comparison’, e.g. the man in John is taller than the man.

The positions on the AH are to be understood as specifying a set of possible grammatical distinctions that a language may make. We are not claiming that any given language necessarily distinguishes all these categories, either in terms of RC formation or in terms of other syntactic processes. For example, some languages (e.g. Hindi) treat objects of comparison like ordinary objects of prepositions or postpositions. In such cases we treat these NPs as ordinary OBLs, and the OCOMP position on the AH is unrealized. Similarly, in Gary and Keenan (1976) it is argued that the DO and IO positions are not syntactically distinguished in Kinyarwanda, a Bantu language. Further, it is possible that in some language RC formation might distinguish between
two types of DOs. If this were so, we would have to expand the AH at that point and say that languages like English do not make the distinction. For the moment, however, we take the AH as specifying the set of possible grammatical distinctions to which RC formation (from simplex main clauses) may be sensitive, since our data do not appear to justify any further refinement in the categories.

In terms of the AH we now give the Hierarchy Constraints (HCs):

**The Hierarchy Constraints (HCs)**

1. A language must be able to relativize subjects.
2. Any RC-forming strategy must apply to a continuous segment of the AH.
3. Strategies that apply at one point of the AH may in principle cease to apply at any lower point.

The HCs define conditions that any grammar of a human language must meet. HC$_1$ says that the grammar must be designed to allow relativization on subjects, the uppermost end of the AH. Thus, for example, no language can relativize only DOs, or only locatives. It is possible, however, for a language to allow relativization only on subjects (and this possibility is in fact realized; see 1.3.1 for examples). HC$_2$ states that, as far as relativization is concerned, a language is free to treat adjacent positions on the AH as the same, but it cannot "skip" positions. Thus, if a given strategy can apply to both subjects and locatives, it can also apply to DOs and IOs. And HC$_3$ states that each point of the AH is a possible cut-off point for any strategy that applies to a higher point. This means that in designing the grammar for a possible human language, once we have given it a strategy that applies at some point on the AH, we are free to terminate its application at any lower point.

Note that it is HC$_3$ that justifies the actual ordering of terms in the AH. Further, HC$_2$ allows as a special case that a particular RC-forming strategy may apply to only a single position. Thus several languages (e.g. Javanese (see Table 1)) have recourse to a case-coding strategy for positions low on the AH (e.g. genitives), whereas the strategy for major NPs is not case-coding. On the other hand, HC$_3$ states that each point on the AH is relevant. Thus, if no language could have an RC-forming strategy that applied to DOs but not to IOs, then the data would not justify making this distinction in the AH. In section 1.3 we substantiate the claim that each point on the AH is in fact a possible cut-off point by showing that for each point on the AH some language has a strategy that cuts off at that point.

1.2.2. The Primary Relativization Constraint. In 1.2.1 we stated that the AH determines, universally, the degree of accessibility to RC formation. But it is perhaps not obvious that the HCs actually make that intuition explicit. However, the Primary Relativization Constraint (PRC) below does make that intuition more explicit, and it logically follows from the HCs. To state it we first define an RC-forming strategy in a
given language to be a primary strategy (in that language) if it can be used to relativize subjects. The term primary is justified in that, by HC$_1$, a language must have a primary strategy but need have no other. Thus, of the various RC-forming strategies a language may have, only primary ones are necessary. We can now state:

The Primary Relativization Constraint (PRC)

1. A language must have a primary RC-forming strategy.
2. If a primary strategy in a given language can apply to a low position on the AH, then it can apply to all higher positions.
3. A primary strategy may cut off at any point on the AH.

Clearly, PRC$_1$ is just a restatement of HC$_1$. PRC$_2$ follows directly from HC$_2$ and the definition of primary, since a primary strategy that can relativize a low position on the AH is one that can relativize subjects (since it is primary) and a low position, and so by HC$_2$ it can relativize all intermediate positions. And PRC$_3$ is simply a special case of HC$_3$.

Thus the PRC states that, restricting our attention to primary RC-forming strategies, relativizability of a low position on the AH entails relativizability of all higher positions. And the converse fails. We may be able to relativize only subjects, both subjects and DOs, or subjects, DOs, and IOs, etc. So the possibility of relativizing (with a primary strategy) decreases as we go down the AH, and in that sense the further we descend the AH the harder it is to relativize.

Note that there is a more obvious, and stronger, way to express the intuition that relativizability decreases as we descend the AH. Namely, “if a language can relativize any position low on the AH, then it can relativize all higher positions”. This claim says in effect that RC-forming strategies always distribute themselves so as to cover an initial segment of the AH. But this formulation is falsified by our data.

Thus Toba Batak (a Malayo-Polynesian language spoken in Sumatra) has a primary strategy, illustrated in (7), which is postnominal and –case:

(7) a. manussi abit boru-boru i.
  wash (active) clothes woman the
  ‘The woman is washing clothes.’

b. boru-boru na manussi abit i
  woman that wash clothes the
  ‘the woman who is washing clothes’

(Note that the basic word order is VOS. See Silitonga (1973) for further substantiation of this point.)

However, direct objects cannot be relativized using this or any other strategy in Toba, as we see from (8a). The only way to achieve the semantic effect of (8a) is first to passivize the underlying sentence, (8b), and then relativize on the derived subject,
1.3.1. Justification of the Hierarchy Constraints

1.3.1. Subjects Only. In many Western Malayo-Polynesian languages, only subjects can be relativized; consider for instance Malagasy, which has basic word order V O X S and a developed system for promoting any major NP to subject position. RCs place
the head NP to the left, followed optionally by an invariable relativizer izay, followed by the restricting clause with no pronoun in the NP\textsubscript{rel} position (Keenan (1972b)):

\begin{enumerate}
\item[a.] Nahita ny vehivavy ny mpianatra.
\hspace*{1em} saw \hspace*{1em} the woman \hspace*{1em} the student
\hspace*{1em} ‘The student saw the woman.’
\item[b.] ny mpianatra izay nahita ny vehivavy
\hspace*{1em} the student \hspace*{1em} that saw \hspace*{1em} the woman
\hspace*{1em} ‘the student that saw the woman’
\item[c.] *ny vehivavy izay nahita ny mpianatra
\hspace*{1em} the woman \hspace*{1em} that saw \hspace*{1em} the student
\hspace*{1em} ‘the woman that the student saw’
\item[d.] Nohitan’ ny mpianatra ny vehivavy.
\hspace*{1em} seen \hspace*{1em} (passive) \hspace*{1em} the student \hspace*{1em} the woman
\hspace*{1em} ‘The woman was seen by the student.’
\item[e.] ny vehivavy izay nohitan’ny mpianatra
\hspace*{1em} the woman \hspace*{1em} that seen \hspace*{1em} the student
\hspace*{1em} ‘the woman that was seen by the student’
\end{enumerate}

Other Malayo-Polynesian languages in our sample that have primary strategies that apply only to subjects are Javanese, Iban, Minang-Kabau, and Toba Batak; also Tagalog, on the assumption that the “focus” NP is the subject (see further 1.4.1.1). Finally, many European languages (e.g. German, Russian, and Polish) have participial RC-forming strategies that apply only to subjects; cf. (2) above.

1.3.2. Subjects—Direct Objects. In Welsh, the primary strategy applies only to subjects and direct objects; it presents postnominal RCs introduced by the particle a, with deletion of NP\textsubscript{rel}. Other positions are relativized by a different postnominal strategy, introduced by the particle y and with a personal pronoun in NP\textsubscript{rel}:

\begin{enumerate}
\item[a.] y bachgen a oedd yn darllen
\hspace*{1em} the boy \hspace*{1em} who was \hspace*{1em} a’ reading
\hspace*{1em} ‘the boy who was reading’
\item[b.] Dyma ’r llyfr y darllenais y stori ynddo.
\hspace*{1em} here-is the book that I-read \hspace*{1em} the story in-it
\hspace*{1em} ‘Here is the book in which I read the story.’
\end{enumerate}

As a second case, one of the primary strategies in Finnish applies only to subjects and direct objects. This strategy places the relative clause prenominally, uses no relativization marker, and puts the subordinate verb in a nonfinite form (different depending on whether the head NP functions as its subject or object). As this strategy is explicitly discussed in Karlsson (1972), we merely note here the illustrative
examples:

(12) a. Pöydällä tanssinut poika oli sairas.
    on-table having-danced boy was sick
    ‘The boy who had danced on the table was sick.’

b. Näkemäni poika tanssi pöydällä.
    I-having-seen boy danced on-table
    ‘The boy that I saw danced on the table.’

As a final case in this category, consider Malay. The primary strategy uses postnominal restricting clauses introduced by the invariable particle *yang* with no pronoun retained in the position relativized:

(13) Ali bunoh ayam yang Aminah sedang memakan.
    Ali kill chicken that Aminah Prog eat
    ‘Ali killed the chicken that Aminah is eating.’

But this strategy applies only to subjects and direct objects. Thus from (14a) we cannot form (14b) or (14c), either stranding or moving the preposition:

(14) a. Ali beri ubi kentang itu kapada perempuan itu.
    Ali give potato the to woman the
    ‘Ali gave the potato to the woman.’

b. *perempuan yang Ali beri ubi kentang itu kapada
    woman that Ali give potato the to

c. *perempuan kapada yang Ali beri ubi kentang itu
    woman to who Ali give potato that

The only systematically elicitable possibility was (15):

(15) perempuan kapada siapa Ali beri ubi kentang itu
    woman to who Ali give potato the
    ‘the woman to whom Ali gave the potato’

Here, the interrogative pronoun *siapa* preceded by the preposition is used to introduce the relative clause. Clearly, the use of the interrogative pronoun, which can take prepositions (i.e. +case), is a different strategy from that using the invariable particle *yang*. In fact, this use of the interrogative pronoun was elicited only under pressure, and was felt to be bookish and clumsy. The preferred alternative was (16), in which the original sentence has been reorganized lexically so that the position relativized is the subject:

(16) perempuan yang menerima ubi kentang itu daripada Ali
    woman that received potato the from Ali
    ‘the woman that received the potato from Ali’

---

2 Our two informants were from Malaya. McDonald and Soenjono (1967) present a slightly more restricted strategy for Indonesian.
1.3.3. Subject—Indirect Object. The indirect object position is perhaps the most subtle one on the AH. For purposes of relative clause formation, it appears that many languages either assimilate indirect objects to the other oblique cases (e.g. English, Malay) or to direct objects (e.g. Shona, Luganda). Nonetheless, Basque does appear to discriminate indirect objects from both its immediate neighbors on the AH.

In Basque, subject, direct object, and indirect object are represented in the verb by verbal affixes (Lafitte (1962, 193–194)). Relativization on any of these positions is effected naturally by deleting the position relativized, putting the restricting clause in prenominal position, and marking the juncture with an invariable marker –n. Thus, from (17a) we can generate any of the three relative clauses (17b)–(17d):

(17) a. Gizon-a -k emakume-a -ri liburu-a eman dio. man the SU woman the IO book the give has ‘The man has given the book to the woman.’
   b. emakume-a -ri liburu-a eman dio-n gizon-a woman the IO book the give has Rel man the ‘the man who has given the book to the woman’
   c. gizon-a -k emakume-a -ri eman dio-n liburu-a man the SU woman the IO give has Rel book the ‘the book that the man has given to the woman’
   d. gizon-a -k liburu-a eman dio-n emakume-a man the SU book the give has Rel woman the ‘the woman that the man has given the book to’

However, once we attempt to relativize on positions that are not explicitly coded in the verb, a variety of difficulties arise. For speakers of what de Rijk (1972) has called the restricted dialect, no further relativization is possible. For other speakers a somewhat greater variety of positions may be relativized, but often a different strategy is used. For instance, a pronoun may be retained in the position relativized and the relative clause may occur postnominally. Consequently, for speakers of the restricted dialect, as well as for certain other speakers, the primary strategy works only on the subject, direct object, and indirect object.

Another language in this category is Tamil, a Dravidian language of southern India and Sri Lanka (Ceylon). One RC-forming strategy puts the restricting clause in prenominal position, with the participial (nonfinite) ending -a on its verb, and no indication in the restricting clause of the syntactic function of NP\(_{rel}\). This strategy applies to subjects, direct objects, and indirect objects:

b. Anta manitan aṭi-tt -a peṇmaṇi(y)-ai jān kaṇ-ṭ āṇ.
    that man hit Past Part woman DO John see Past Sg-3rd-Masc
    'John saw the woman that that man hit.'

c. Jān puttetak-ass(a) koṭi-tt -a peṇmaṇi(y)-ai nān kaṇ-ṭ āṇ.
    John book DO give Past Part woman DO I see Past Sg-1st
    'I saw the woman to whom John gave the book.'

It does not work on other positions, e.g. instrumentals, where a different RC-forming strategy is required, retaining NP_rel in the restricting clause:

(19) Eṇna(k) katti(y)-āḷ koṭi(y)-ai anta manitan
    which knife with chicken DO that man
    kolaippe-tt āṇ anta katti(y)-ai jān kaṇ-ṭ āṇ.
    kill Past Sg-3rd-Masc that knife DO John see Past Sg-3rd-Masc
    'John saw the knife with which the man killed the chicken' (literally: with which knife the man killed the chicken, John saw that knife).

Roviana, a Melanesian language spoken in New Georgia, Solomon Islands, also provides some support for discriminating the indirect object position. There, in simplex sentences, indirect objects are treated like other oblique case NPs in that they are preceded by a preposition, whereas direct objects are not:

(20) Ele ponia Jone koe Mere sa buka.
    Past give John to Mary the book
    'John gave the book to Mary.'

(21) Vekoa Jone sa bereti pa tevelo.
    put John the bread at table
    'John put the bread on (the) table.'

On the other hand, in relative clauses indirect objects are relativized just like direct objects and subjects—a postnominal strategy, in which the case of the relativized position is not marked, as the relativization marker sapu is morphologically invariable and no pronoun is retained in the position relativized. But in the oblique cases the function of NP_rel is coded, either in the form of a stranded adverb or in the variable form of the relativization marker:

(22) a. sa buka sapu ele ponia Jone koe Mere
    the book that Past give John to Mary
    'the book that John gave to Mary'

    b. sa barikalege sapu ele ponia buka Jone
    the woman that Past give book John
    'the woman that John gave the book to'

(23) sa tevelo vasina vekonia Jone sa bereti
    the table where put (by?) John the bread
    'the table where John put the bread'
1.3.4. **Subject—Oblique.** The primary strategy in Korean (Tagashira (1972)) places the restricting clause to the left of the head NP, separated from it by the suffix -(i)n, -nin. NP_rel is simply deleted for all NPs on the AH down to and including obliques:

(24) hyønsik -i ki lā -lil ttāli-n maktāki
Hyensik SU the dog DO beat Rel stick
‘the stick with which Hyensik beat the dog’

However, where genitives are relativized, a pronoun must be retained:

(25) chaki-ij lā -ka chongmyøngha-n ki salam
he of dog SU smart Rel the man
‘the man whose dog is smart’

In fact, many languages besides Korean change RC-forming strategies at the genitive position by presenting a pronominal element in the position relativized; see Table 2 in section 2.2.2.

Further support for distinguishing genitives on the AH comes from languages like Catalan and North Frisian (Fering dialect), where genitives (and objects of comparison) are simply not relativizable at all, although all NPs higher on the AH than genitives are relativizable, e.g. North Frisian:

(26) a. John kland det wuf’s henk.
   ‘John stole the woman’s chicken.’
   b. det henk wat kland John
   ‘the chicken that John stole’
   c. *det wuf wat’s henk John kland
   ‘the woman whose chicken John stole’

1.3.5. **Subject—Genitive.** A great many well-known languages have primary strategies that permit relativization on all the positions on the AH except that of objects of comparison. Thus in French we cannot relativize le jeune homme in (27a):

(27) a. Marie est plus grande que le jeune homme.
   ‘Marie is bigger than the young man.’
   b. *le jeune homme que que Marie est plus grande
   ‘the young man than whom Marie is bigger’

The same situation holds in Spanish, German, and Romanian (nonparticipial strategy, in each case).

1.3.6. **Subject—Object of Comparison.** Few languages that distinguish objects of comparison from direct objects or oblique NPs permit them to be relativized. In English we do have phrases like (28), though some find them rather uncomfortable:

(28) the man who Mary is taller than
And in Urhobo, whose primary strategy is postnominal, and pronoun-retaining for all positions on the AH, we have:

(29) oshale na l- i Mary rho n- o
     man the that Mary big than him
     'the man that Mary is bigger than'

This completes the argument that primary RC-forming strategies can discriminate all the positions on the AH. Thus, the PRC is established, and we may conclude generally that the Accessibility Hierarchy determines the relative ease of relative clause formation from unmarked simplex sentences across languages.

Table 1 (p. 76) summarizes our data concerning the relativizing power of the RC-forming strategies we considered in the languages in our sample. A key to the entries in the row–column intersections is given at the end of the table. Needless to say, such a summary cannot adequately represent all the language-particular problems involved in determining the nature of an RC-forming strategy.

1.4. Problems and Possible Counterexamples

We shall consider here two types of difficulties with the analysis we have proposed. The first concern methodological problems involved in identifying subjects and relative clauses (RCs). The second concern specific counterexamples to the Hierarchy Constraints.

1.4.1. Methodological Problems

1.4.1.1. Identifying Subjects. We are using a largely traditional notion of subject. An attempt to make explicit the large number of specific properties that comprise this notion is given in Keenan (1976b) and will not be discussed here. It is clear from that investigation, however, that the NPs we call subjects in some languages are more subject-like than those of other languages. That is, they possess a greater number of the properties that are characteristic of subjects in general. A language in which the subject properties were systematically distributed across two or more NPs then might arguably be said not to have a single category of subject. In such a case the AH for that language would lack the subject position, much as the AH applied to other languages may lack an OCOMP position (see 1.2.1), and the predictions made by the HCs would be reduced and further principles would have to be found to account for the degree of relativizability of NPs not on the AH.

Fortunately, many languages do appear to present NPs that conform fairly well to the traditional concept of subject. Nonetheless, three categories of possibly subjectless languages have been recently discussed in the literature: Tagalog and Philippine languages generally (Schachter (1976)), "topic" oriented languages like Lisu and other Sino-Tibetan languages (Li and Thompson (1974; 1976)), and ergative languages (Tchekhoff (1973)) (see 1.4.2 below).
Table 1

<table>
<thead>
<tr>
<th>Language</th>
<th>Relative clause forming strategy</th>
<th>Relativizable positions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Subj</td>
</tr>
<tr>
<td>Aoban (North-East)</td>
<td>1. postnom, –case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>–</td>
</tr>
<tr>
<td>Arabic (Classical)</td>
<td>1. postnom, –case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>–</td>
</tr>
<tr>
<td>Basque</td>
<td>1. prenom, –case</td>
<td>+</td>
</tr>
<tr>
<td>Batak (Toba)</td>
<td>1. postnom, –case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>–</td>
</tr>
<tr>
<td>Catalan</td>
<td>1. postnom, –case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>–</td>
</tr>
<tr>
<td>Chinese (spoken Pekingese)</td>
<td>1. prenom, –case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. prenom, +case</td>
<td>–</td>
</tr>
<tr>
<td>Czech (colloquial)</td>
<td>1. postnom, +case</td>
<td>+</td>
</tr>
<tr>
<td>Dutch</td>
<td>1. postnom, –case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>–</td>
</tr>
<tr>
<td>English</td>
<td>1. postnom, –case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>–</td>
</tr>
<tr>
<td>Finnish</td>
<td>1. postnom, +case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. prenom, –case</td>
<td>+</td>
</tr>
<tr>
<td>French</td>
<td>1. postnom, +case</td>
<td>+</td>
</tr>
<tr>
<td>North Frisian (Fering dialect)</td>
<td>1. postnom, –case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>–</td>
</tr>
<tr>
<td>Fulani (Gombe dialect)</td>
<td>1. postnom, –case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>–</td>
</tr>
</tbody>
</table>
Table 1 (Continued)

<table>
<thead>
<tr>
<th>Language</th>
<th>Relative clause forming strategy</th>
<th>Subj</th>
<th>DObj</th>
<th>IObj</th>
<th>Obl</th>
<th>Gen</th>
<th>OComp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genoese (Zeneyze)</td>
<td>1. postnom, -case</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>German</td>
<td>1. postnom, +case</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2. prenom, -case</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gilbertese</td>
<td>1. postnom, -case</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>*</td>
</tr>
<tr>
<td>Greek (Modern)</td>
<td>1. postnom, -case</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Hausa</td>
<td>1. postnom, -case</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Hebrew</td>
<td>1. postnom, -case</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-/*</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+/*</td>
</tr>
<tr>
<td>Hindi</td>
<td>1. postnom, +case</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>2. internal, +case</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>*</td>
</tr>
<tr>
<td>Iban (Sea Dayak)</td>
<td>1. postnom, -case</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+/*</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Italian</td>
<td>1. postnom, -case</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-/*</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-/*</td>
</tr>
<tr>
<td>Japanese</td>
<td>1. prenom, -case</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+/*</td>
<td>+/*</td>
<td>-?</td>
</tr>
<tr>
<td></td>
<td>2. prenom, +case</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+/*</td>
<td>-</td>
<td>-?</td>
</tr>
<tr>
<td>Javanese</td>
<td>1. postnom, -case</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Kera</td>
<td>1. postnom, -case</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>*</td>
</tr>
</tbody>
</table>
Table 1 (Continued)

<table>
<thead>
<tr>
<th>Language</th>
<th>Relative clause forming strategy</th>
<th>Relativizable positions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Subj</td>
</tr>
<tr>
<td>Korean</td>
<td>1. prenom, -case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. prenom, + case</td>
<td>-</td>
</tr>
<tr>
<td>Luganda</td>
<td>1. postnom, + case</td>
<td>+</td>
</tr>
<tr>
<td>Malagasy</td>
<td>1. postnom, -case</td>
<td>+</td>
</tr>
<tr>
<td>Malay</td>
<td>1. postnom, -case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. postnom, + case</td>
<td>-</td>
</tr>
<tr>
<td>Maori</td>
<td>1. postnom, - case</td>
<td>+</td>
</tr>
<tr>
<td>Minang-Kabau</td>
<td>1. postnom, - case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. postnom, + case</td>
<td>-</td>
</tr>
<tr>
<td>Persian</td>
<td>1. postnom, - case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. postnom, + case</td>
<td>-</td>
</tr>
<tr>
<td>Polish</td>
<td>1. postnom, + case</td>
<td>+</td>
</tr>
<tr>
<td>Romanian</td>
<td>1. postnom, + case</td>
<td>+</td>
</tr>
<tr>
<td>Roviana</td>
<td>1. postnom, - case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. postnom, + case</td>
<td>-</td>
</tr>
<tr>
<td>Russian</td>
<td>1. postnom, + case</td>
<td>+</td>
</tr>
<tr>
<td>Sinhala</td>
<td>1. prenom, - case</td>
<td>+</td>
</tr>
<tr>
<td>Shona</td>
<td>1. postnom, - case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. postnom, + case</td>
<td>-</td>
</tr>
<tr>
<td>Slovenian</td>
<td>1. postnom, + case</td>
<td>+</td>
</tr>
</tbody>
</table>
Table 1  (Continued)

<table>
<thead>
<tr>
<th>Language</th>
<th>Relative clause forming strategy</th>
<th>Relativizable positions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Subj</td>
</tr>
<tr>
<td>Spanish</td>
<td>1. postnom, -case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>-</td>
</tr>
<tr>
<td>Swedish</td>
<td>1. postnom, +case</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2. postnom, -case</td>
<td>+</td>
</tr>
<tr>
<td>Tagalog</td>
<td>1. postnom, -case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. prenom, -case</td>
<td>+</td>
</tr>
<tr>
<td>Tamil</td>
<td>1. prenom, -case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. internal, +case</td>
<td>+</td>
</tr>
<tr>
<td>Tongan</td>
<td>1. postnom, -case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>+</td>
</tr>
<tr>
<td>Turkish</td>
<td>1. prenom, -case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. prenom, +case</td>
<td>-</td>
</tr>
<tr>
<td>Urhobo</td>
<td>1. postnom, +case</td>
<td>+</td>
</tr>
<tr>
<td>Welsh</td>
<td>1. postnom, -case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>-</td>
</tr>
<tr>
<td>Yoruba</td>
<td>1. postnom, -case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>-</td>
</tr>
<tr>
<td>Zurich German</td>
<td>1. postnom, -case</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>2. postnom, +case</td>
<td>-</td>
</tr>
</tbody>
</table>

Key: + means that the strategy generally applies to that NP position; − means that it does not. +? means it applies, but with loss of acceptability and perhaps some informant disagreement. −? means it does not apply, although the result is not judged too bad by informants. An entry of the form x/y means that for certain of the NPs in that position x is the correct entry, and for other such NPs y is the correct entry. x(?) means that our data are not entirely unequivocal, but our best judgment is that x is the appropriate entry. * means that that NP position does not exist as such, but rather is treated as some other position: e.g. objects of comparison are often either direct objects of verbs like exceed as in Shona or oblique case NPs governed by pre- or postpositions as in Hebrew and Hindi. A blank means that we lack the relevant data.
Of these categories, Schachter's claim that the "focus" NP in Tagalog cannot be regarded as a subject is the most damaging to our claim, since Tagalog was one of the languages that justified HC$_1$ and HC$_3$ (a language may relativize only subjects). However, it has been shown in Keenan (1976a) that others of the Western Malayo-Polynesian languages, notably Malagasy, do not present the evidence that supports Schachter's claims for Tagalog, and consequently there is still sufficient data to support HC$_1$.

The evidence Li and Thompson present from Sino-Tibetan is less damaging, since the NPs that are most subject-like do not present any Hierarchy violations. What they argue, however, is that the syntactic category "subject" has a very low functional load in these languages in that few if any syntactic processes need to be made sensitive to the NP that is the subject. Further research in those languages might then reveal that categories other than those on the AH, e.g. topic, are relevant in determining RC-forming possibilities.

With the possible exception of Dyirbal (and perhaps Eskimo; see Woodbury (1975)) (see section 1.4.2.1), the claim that ergative languages lack subjects has been refuted by Anderson (1976), who shows that with respect to many major syntactic processes transitive subjects (the ergative NP) and intransitive subjects behave alike, in distinction to DOs, thus establishing the fact that there is a category of subject in those languages.

1.4.1.2. Identifying Relative Clauses. We have been considering an RC to be any syntactic structure that designates an object (or set of objects) in a certain way, namely, by first specifying a larger domain of objects and then restricting it to a subset, perhaps a one-member subset, of which a certain sentence, the restricting sentence, is true. However, many languages present sentence types that appear to designate objects in this way but in which there is no surface constituent with which we can associate the designating properties. One such example concerns extraposed RCs, as in (30a) below:

(30) a. The student finally arrived who we had been waiting all morning for.
    b. the student who we had been waiting all morning for

Clearly, to evaluate the truth of (30a) it is necessary to determine that the main predicate holds of an object that is, first, a student, and second, is such that the sentence \textit{we had been waiting all morning for him} is true of him. That is, the object in question is clearly the one designated by the RC in (30b), and in this sense the logical structure of (30a) contains the RC in (30b). But (30a) itself does not present in surface a constituent that has the referential properties of (30b) and so does not contain an RC.

In this case, however, it is easy to argue that at a fairly shallow level of underlying structure (30b) does occur in the syntactic structure of (30a), and that a late rule of RC Extraposition moves the restricting clause around the main verb. Thus the syntactic structure of (30a), considered as the sequence of phrase markers representing its
derivation, does contain an RC, and it is clear, then, that our definition of RC does not require that, in surface, the head NP and the restricting clause be a constituent.

It is clear, furthermore, that the most ordinary type of RC in English is expressed by structures like (30b), those in (30a) being more marked and of more restricted distribution. But in many languages the usual translations of English sentences containing RCs do not present the head NP and the restricting clause as a constituent. (31) from Hindi and (32) from Walbiri (Hale (n.d.)) illustrate this type:

(31) Mai us aurat ko janta hoon jis- ko Ram ne kitab diya. 
      | that woman DO know who IO Ram Erg book gave 
      ‘I know the woman that Ram gave the book to.’

(32) नात्जुल-जु φ-νा yankiri pantu-νु, kutja-lpa नापा ना- नु. 
      | Erg Aux emu spear Past CompAux water drink Past 
      ‘I speared the emu that was drinking water/while it was drinking water.’

In these cases it is not clear that what corresponds to the restricting clause in English ever occurs as a constituent with the head NP in underlying structure. Both Hale (n.d.) and Andrews (1975) argue that it does not.

However, languages with this type of construction very regularly exhibit a related type of construction in which there is a constituent that meets our semantic conditions of relative clause-hood. The corresponding versions of the Hindi ((31)) and Walbiri ((32)) are given below in (33) and (34), respectively:

(33) [Ram ne jis aurat ko kitab diya] us (aurat) ko mai janta hoon. 
      Ram Erg which woman IO book gave that woman DO I know 
      ‘I know the woman that Ram gave the book to.’

(34) [yankiri-लु kutja-lpa नापा ना- नु] ula pantu-νु नात्जुल-जु. 
      emu Erg CompAux water drink past that-one spear Past I Erg 
      ‘I speared the emu that was drinking water.’

It is not clear that the bracketed constituents above should be considered NPs. Certainly they present the normal syntax of full sentences, except that they contain the markers jis and kutja. Whatever the grammatical category of these constituents, however, it does appear that they meet our definition of RC—they specify a domain of objects, those marked by jis and kutja, and they give the restricting clause, the one determined by the entire clause in which they occur. Thus, like the examples (3) from Bambara and (4) from Digueño mentioned earlier, they illustrate an RC-forming strategy in which the head NP occurs within the restricting clause. Our definition of RC then does not require that an RC be an NP, nor does it require that the head NP command the restricting clause (a use of “head” that differs from the usual one in the linguistic literature).

The RC-forming strategies illustrated above in Walbiri are perhaps responsible for the claim sometimes heard that, in violation of HC1 (which states that subjects are
universally relativizable), Australian languages do not have RCs. Walbiri at least clearly does not present embedded RCs in which the head NP commands the restricting clause and in which the two form an NP constituent. But it certainly does present constituents that meet our semantically-based criteria for what constitutes an RC, so HC1 is not violated here.

1.4.2. Some Possible Counterexamples

1.4.2.1. Possible Counterevidence to HC1. We consider first the best documented counterexample to HC1 (subjects are always relativizable). The counterevidence comes from Dyirbal (Dixon (1969; 1972)), an Australian language. Dyirbal appears ergative in the standard sense that full NP subjects of intransitive sentences and full NP DOs of transitive sentences are case-marked in the same way (zero, in this case), whereas subjects of transitive sentences carry a special marker, the ergative. In the analysis to follow we shall refer to intransitive subjects and transitive DOs collectively as absolutives. Examples are from Dixon (1972, 100–101).

(35) bayi yara banagajiu.
    Det-Abs man-Abs return
    ‘The man is returning.’

(36) bayi yuri bangul yarangu bagan.
    Det-Abs kangaroo-Abs Det-Erg man-Erg spear
    ‘The man speared a kangaroo.’

Dixon demonstrates that absolutives can be relativized from both transitive and intransitive sentences, but the ergative NP cannot be relativized. To talk about the man who speared a kangaroo, it is necessary to apply a transformation that promotes the ergative NP to absolute status (37a), whence it can be relativized (37b):

(37) a. bayi yaşı bagalnaŋu bagul yurigu.
    Det-Abs man-Abs spear-‘Pass’ Det-Instr kangaroo-Instr
    ‘The man speared the kangaroo.’

b. [Np] bayi yaşı bagal-ŋa -ŋu bagul
    Det-Abs man-Abs spear ‘Pass’ Rel Det-Instr
    yurigu] banagaŋu.
    kangaroo-Instr return
    ‘The man who speared the kangaroo is returning.’

On the basis of the Dyirbal data, Johnson (1974a) has suggested that the upper segment of the AH is inapplicable to ergative languages generally, and that an Ergative Hierarchy (EH), ABS > ERG > IO > OBL, etc. more adequately describes the accessibility to RC formation in those languages. The EH differs significantly from the AH only in that it predicts the possibility of RC-forming strategies that relativize only absolutives, such as in Dyirbal. But general support from ergative languages for this prediction is lacking. The majority of ergative languages known to us permit both
ergative and absolutive NPs to relativize. We have already illustrated relativization on ergatives in Basque (17b) and Walbiri (34). Examples (38)–(42) below illustrate relativization on ergatives from a few other languages.

**Hindi**

(38) Ram us ḥarki ko pasand kartahe, jis- ne kapre dhoye.  
Ram that girl DO likes Rel Erg clothes washed  
‘Ram likes the girl who washed clothes.’

**Tongan (Anderson (1975))**

(39) te mo fetaulaki mo e tangata ’oku ne fua ’a e sioki vai.  
Fut 2Du meet with Art man Pres 3Sg carry Part Art jug water  
‘You will meet a man who is carrying a jug of water.’

**Gahuku (Deibler (1973))**

(40) izegipa get -a ve  
child begot 3Sg man  
‘the man who had begotten the child’

**Jacaltec (Craig (1974))**

(41) x- φ- w- ’il [naj x- φ- watxe-n hun ti’]  
Asp him-Abs I-Erg see man Asp him-Abs make Rel one this  
‘I saw the man who made this.’

**Greenlandic Eskimo (Woodbury (1975))**

(42) qimmi-φ -a tuquk-kiga nalu Nil- saNa.  
dog Abs-Sg Poss-3Sg kill Trans ignorant Neg Indic-3Sg-1Sg  
‘He knows about me, who killed his dog.’

Woodbury (1975) does argue, however, that absolutes are more relativizable in Greenlandic than are ergatives, on the grounds that (1) RCs formed on ergatives are somewhat more restricted in the distribution in matrix clauses (p. 21) than are those formed on absolutes, and (2) for certain verb classes ergatives cannot be relativized out of the active participle (p. 27). To relativize that NP from an active participle, the ergative must first be promoted to an absolute via Antipassive. Nonetheless, it is the case that ergatives can in general be relativized, even if not quite so freely as absolutes, and so the AH does appear to apply, although more work would need to be done to distinguish according to our criteria what the different RC-forming strategies in Greenlandic are.

The general claim, then, that in ergative languages absolutes are more relativizable than ergatives receives little support. This in turn suggests that the Dyirbal data might be analyzed differently. One reanalysis that would have the effect of making Dyirbal conform to the AH would be simply to regard the absolute NP in transitive sentences (as well as in intransitive ones) as the subject. Such an analysis again calls into question the defining criteria for subjecthood and would require much independent
support. Perhaps surprisingly, however, Dyirbal does provide such support. Below we summarize this evidence, citing page numbers from Dixon (1972) for the supporting data.

1. Absolutives are the most essential NP in the sentence. Thus an unspecified agent (ergative) may simply be eliminated from a sentence, much like unspecified agents of passive sentences in English. But unspecified absolutives cannot be so eliminated. If we want to eliminate an unspecified patient, we must first "passivize" the sentence so that the ergative becomes a surface absolutive and the former absolutive becomes oblique, whence it can be deleted when nonspecific (p. 70).

2. Absolutives are the only target of advancement rules. Dyirbal has several rules, rather like Passive in English, that promote e.g. ergatives and instrumentals to the absolutive position of underlying transitive sentences, as we illustrated in (37a).

3. Absolutives normally precede ergatives in simplex sentences (p. 291).

4. Only absolutes can be coreferentially deleted by operations like Equi and Conjunction Reduction (pp. 73, 67). To coreferentially delete an ergative NP, it must first be advanced to an absolutive.

5. It appears that certain demonstratives, always definite and presupposing reference—a characteristic property of subjects—are restricted to absolutive positions (p. 218).

6. In the few cases where we can tell, absolutes appear to control reflexives. Usually reflexivization is done by verbal affixing, so the derived verb is intransitive and the subject is absolutive on independent grounds. However, in a few examples the object is an inalienable body part and is expressed in surface structure even though the verb is reflexivized. And in these cases the controller is absolutive (p. 153):

   (43) bayi yaça mala ḍaŋgaymarĩnu.
   Det-Abs man-Abs hand-Abs eat-Refl-Pres/Past
   'The man chewed his finger.'

If we adopt the analysis of Dyirbal in which absolutes in transitive sentences are the subjects, then Dyirbal will be typologically anomalous in only one major respect: namely, in the least marked type of transitive sentence the NP with the referential or topic properties of subjects does not express the agent. And since the predominance of topic properties over agency properties in this case is quite large (very few transformations mention ergatives, but very many mention absolutes), it turns out that the most subject-like NP in basic transitive sentences is the absolutive and does not express the agent.

This analysis at least has the (metalinguistic) advantage of isolating at a single point the anomaly of Dyirbal—namely, transitive subjects in basic sentences are not agents. The ergative analysis on the other hand makes Dyirbal anomalous in a great many respects: the word order is OSV; DOs are the most essential constituent of transitive sentences, subjects being freely deletable when unspecified; etc. Further-
more, it is possible that when the lexicons of more languages have been analyzed in detail, it will be seen that many languages have classes of verbs that select goal subjects but allow agents present in oblique cases. Biggs (1974) has recently argued, for instance, that Fijian has one large class of verbs that select goal rather than agent subjects, whereas another class selects agent subjects, as is more usual. Perhaps Dyirbal will then be seen to be merely near the end of a continuum in that most of its transitive verbs select nonagent subjects.

1.4.2.2. Apparent Exceptions to \( HC_2 \). A second type of counterexample we consider are apparent exceptions to \( HC_2 \), that a given RC-forming strategy must apply to a continuous segment of the AH. Hausa (Abrahams (1959); Schachter (1973)) and Yoruba appear to use RC strategies that apply to discontinuous segments of the RC.

Thus Hausa seems to use a +case strategy to relativize subjects, since a personal pronoun seems present in the position relativized:

\[
(44) \quad \text{dokin da ya mutu} \\
\quad \text{horse Rel it died} \\
\quad \text{‘the horse that died’}
\]

Relativization on DOs, however, does not permit retention of a pronoun in the position relativized, and as no other nominal particle is present to code the case of the relativized NP (the relativizer \( da \) is invariable), RC formation on DOs appears to be effected by a −case strategy:

\[
(45) \quad \text{mutumin da na gani (*shi)} \\
\quad \text{person Rel I saw him} \\
\quad \text{‘the person that I saw’}
\]

However, when we relativize on OBLs in Hausa, a pronoun is again present. (In relativizing on IOs, the presence of a pronoun is optional.)

\[
(46) \quad \text{yaron da suka gaya {wa/masa}} \\
\quad \text{child Rel they said to to-him} \\
\quad \text{‘the child whom they told’}
\]

\[
(47) \quad \text{wuqad da ya kashe ta da ita} \\
\quad \text{knife Rel he killed her with it} \\
\quad \text{‘the knife with which he killed her’}
\]

It appears, then, that the postnominal, +case strategy in Hausa does not apply to a continuous segment of the AH, in violation of \( HC_2 \).

On examining further the pronoun that appears in the RC when a subject is relativized, however, we note that this same pronoun is also required in simplex
sentences with a subject NP:

(48) a. Yusufu ya zo.
    Joseph he came
    ‘Joseph came.’

b. *Yusufu zo.

Quite generally in Hausa, full subject NPs must be accompanied by clitic pronouns. This suggests a different line of analysis for such pronouns. Instead of regarding them as constituent parts of a +case strategy, we may regard them as an instance of verb agreement. In many languages verb-agreement affixes are known to derive from (clitic) pronouns (Givón (1976)), and it is often difficult to draw a precise dividing-line between clitic pronouns and agreement affixes. We are therefore suggesting that where the presence of a pronoun is required by the presence of a full NP, then the pronoun be regarded as an instance of verb agreement, and not as an instance of NP case marking. In Hausa, such “internal pronouns” (for this terminology, see Keenan (1972a, 447–450)) occur only with subjects; with DOs, etc., we have (49a) and (49c) for instance, but not (49b) and (49d):

(49) a. Na gani mutumin.
    b. *Na gani shi mutumin.
    ‘I saw the person.’

    c. Ya kashe ta da wuqad.
    d. *Ya kashe ta da ita da wuqad.
    ‘He killed her with the knife.’

Another similar possible counterexample, from Tongan (pointed out in Anderson (1975)), is not so easily disposed of, however. Here it appears that pronouns may, and in some cases must, be retained when subjects are relativized. (39) already illustrates this. Similarly, relativization on IOs, OBLs, and GENs obligatorily leaves behind a pronominal trace (examples from Anderson (1975), Chung (personal communication)):

(50) ko e ’eiki eni na’e lango mo’ona ’a e fale lahi.
    Part Art chief this Past build for-3Sg Part Art house big
    ‘This is the chief for whom the big house was built.’

(51) fakahā mai kiata au ’a e tamasi’i na’e show towards-to-Pers me Abs Art child Past
    disappear his brush toothbrush
    ‘Show me the boy whose toothbrush disappeared.’

However, it is not possible to present pronominal traces when DOs are relativized:

(52) ’oku ’ikai ’ilo ’e ha taha ’a e tangata na’a ku taa’i (*ia).
    Pres not know Erg any one Abs Art man Past I hit him
    ‘Nobody knows the man who I hit.’
Furthermore, the type of explanation used above for Hausa and Yoruba does not in general seem applicable. The subject pronouns present when subjects are relativized do not in general function as agreement particles, although the pronoun present in the RC is the clitic one, not the independent one. Thus it occurs preverbally, whereas the independent pronouns and full NPs functioning as subjects normally occur postverbally (Tongan is VSO).

(53) na’e ’ave ’e Sione ho’o telefone.
    Past take Erg John your telephone
    ‘John took away your telephone.’

(54) *na’a ne ’ave ’e Sione ho’o telefone.
    Past he take Erg John your telephone

Furthermore, the clitic pronouns can be present when the subject is an independent pronoun, although the meaning here is emphatic:

(55) na’a ne taa’i ’e ia pē ia.
    Past he hit Erg he Emp him
    ‘He hit himself.’

This use of the clitic could conceivably develop so that the clitic would cooccur with full NPs in subject position without the emphatic meaning. This at least would be a normal way for an agreement to arise (see Givón (1976) for justification of this claim). But this would be to anticipate the development of Tongan and cannot be used to justify its current status as a counterexample to HC₂.

Perhaps the best explanation we can offer for the pattern of RC-forming strategies in Tongan is a historical one. As we have mentioned, Tongan is an ergative language in which the ergative marker ’e is cognate with the passive agent marker in related Polynesian languages. Hohepa (1969) has presented evidence that the ergative paradigm present in Tongan and the closely related Niuean has evolved from an original nominative–accusative paradigm in which the productive distinction between active and passive verb forms was lost and the passive morphology retained on the major NPs. Thus, historically speaking, the ergative subject of Tongan is a passive agent, and the absolutive DO in Tongan is historically the subject. We suggest here that the evolution of the agent NP to subject status in Tongan is not quite complete. While Anderson (1976) has shown that it does possess many general properties of subjects, it may retain a few traces of its original oblique case status. For example, in simple active sentences the ergative NP can be omitted when unspecified, like the ergative NP in Dyirbal and the passive agent in English, as illustrated in (56) and (57) below (both from Churchward (1953)):

(56) na’e tamate’i ’e Tevita ’a Koliate.
    Past kill Erg David Abs Goliath
    ‘David killed Goliath.’
(57) na’e tamate’i ‘a Koliate.
Past kill Abs Goliath
‘Goliath was killed.’

Now, since OBLs in general in Tongan require retention of a clitic pronoun when relativized (in one case a nonclitic pronoun can be retained, but this is quite exceptional), the ergative NP in Tongan is behaving like the other OBLs in this respect. On the other hand, the DO does not allow a pronoun to be retained under relativization. But, as we show in 2.2.2, this is characteristic of subjects in general. That is, even in languages like Hebrew in which pronoun retention is normal under relativization, subject pronouns are normally not retained. So, in this way, the DO in Tongan is behaving like a subject. In these very minor respects, then, perhaps the ergative and absolutive in Tongan betray their earlier status as passive agent and subject, respectively.

2. Towards an Explanation of the Hierarchy Constraints

2.1. The AH as a Psychologically Valid Entity

Given that the HCs do make correct predictions about RC formation in a wide range of languages, it is natural to wonder why this should be so. We propose the following explanation: The AH directly reflects the psychological ease of comprehension. That is, the lower a position is on the AH, the harder it is to understand RCs formed on that position.

If the AH does reflect the psychological accessibility of NPs to RC formation, then we can use this fact, in conjunction with certain other assumptions, to explain the HCs in the following way. First, it would be natural that a way of relativizing a certain position might not be applicable at the next lower position (HC3) on the general assumption that syntactic processes are ways of encoding meanings; and, if one meaning is inherently more difficult to encode than another, then a strategy for encoding the first need not apply to the second. By the same token, a strategy that applies to one position but fails to apply to the next lower position would not be expected to apply to a still lower position (HC2). For, if a given strategy is used to encode a fairly easy meaning and that strategy is “strong” enough to encode a rather difficult meaning, then it is surely strong enough to encode the meanings of intermediate difficulty. However, this rather informal line of reasoning can be slightly extended to yield an argument that languages should not have “gapped” RC-forming strategies. Thus, if speakers can in general encode a meaning of a certain difficulty in some way, then they should certainly be able to encode meanings of lesser difficulty (possibly in a different way). Yet we have seen that Toba Batak, for example, can relativize directly subjects using one strategy, and IOs, OBLs, and GENs using another. But DOs cannot be directly relativized using any strategy. As we mentioned, however, in all cases in our sample in which languages present gapped strategies, the NPs that occur in
the gap can always be promoted to positions that can be directly relativized via operations like Passive. In these cases, then, we would like some independent evidence that it is psychologically easier, in those languages, to for instance promote a DO to subject and then relativize it rather than to relativize it as a DO. We have no direct psychological evidence of the right sort. It is worth noting, however, that in those languages that promote an NP to a higher position on the AH in order to relativize it, the promotion system (e.g. the ways of converting NPs low on the AH to ones higher on the AH) is usually well developed, very commonly used, and has a wide syntactic distribution, in distinction for instance to Passive in English, which is, by comparison, a less usual, more marked form. In fact, in Hawkins and Keenan (1974) it is reported that on certain types of repetition tasks, English-speaking children (10–12 years) do significantly less well on RCs like the boy who was seen by Mary than they do on ones like the boy who Mary saw. Thus, promoting to relativize is certainly not universally easier than relativizing directly.

Finally, note that this psychological interpretation of the AH cannot fully account for HC1, that subjects must, in general, be relativizable. It only justifies the claim that subjects are easier to relativize than any other position on the AH, but it would allow in principle that, in some language, no position on the AH be relativizable.

2.1. Evidence in Support of the Psychological Validity of the Accessibility Hierarchy

Several recent experimental studies do provide partial support for the claim that the AH does represent the psychological accessibility to RC formation. Thus, in Legum (1975) it is shown that English-speaking children, aged 6–8, comprehend RCs formed on subjects better than ones formed on DOs. Similar results were achieved by Brown (1971) for children aged 3–5 years. Hatch (1971) shows that young children respond more rapidly to RCs formed on subjects than on objects, and Cook (forthcoming) has shown that children and adults recognize with fewer errors instances of head NPs bearing the subject relation to the subordinate verb in RCs than instances of the head bearing the object relation to the verb. And Valli et al. (1972), in a production test, show that French children at the level of 6e (approx. 12 years of age) produce RCs formed on subjects much more readily than on objects, and RCs on possessor NPs are much rarer still and in the majority of instances incorrectly formed.

It should be noted that the purpose of these studies was not specifically to validate the AH. Variables other than the position on the AH of the NP relativized were in general shown to be relevant (e.g. whether the RC was embedded or not, whether the function of the RC in the matrix was the same as that of the head in the RC (Sheldon (1974)); see Legum (1975) for some discussion of these other parameters).

On the other hand, Hawkins and Keenan (1974) report a study specifically designed to test the psychological validity of the AH. Children, aged 10–12, were given repetition tests involving RCs formed on all the positions on the AH. Correctness of recall correlated significantly (<.05) with the AH. In fact, the recall hierarchy
established in that study was:

\[ SU > DO > IO > OBL, OCOMP > GEN \]

Thus, RCs formed on subjects were recalled with fewer errors than those formed on DOs, which were better than ones formed on IOs, which were better than those formed on either OBLs or OCOMPs, which were better than ones formed on possessor NPs. Clearly the only position out of place on the recall hierarchy is the OCOMP position, which was treated by the children as the same in accessibility as the OBL position. All RCs on OBLs left the preposition stranded, e.g. the boy who Johnny took the toy from. Apparently stranding the comparative particle, e.g. the boy who Johnny is taller than, was interpreted as being similar to preposition stranding.

While it would certainly be premature, on the basis of these few studies, to conclude that the AH does establish a hierarchy of psychological accessibility, it does appear that the available evidence points in that direction. If further research along these lines justifies the psychological interpretation of the AH, then we will have an explanation for two further facts from our own study. Alternatively, these facts can be considered further support for the psychological interpretation of the AH. The first concerns a pattern of judgments of relative accessibility of RCs from languages in which most positions on the AH can be grammatically relativized. The second concerns the distribution of personal pronouns retained in the RC.

2.2.1. Intralanguage Relative Accessibility Judgments. One natural way to extend the HCs would be to interpret them not only as a cross-language ordering of grammaticality, as we have done, but to consider them as an acceptability ordering within each language. Thus one might expect that in general, within a given language, RCs formed on the high end of the AH would be judged more acceptable than RCs formed on the low end. And in the extreme cases this appears to be correct.

RCs formed on subjects are always among the most acceptable in any given language, and those formed on objects of comparison, where possible at all, are often judged to be only marginally acceptable. For instance, many informants in English are uncomfortable with the man who Mary is taller than. Similar judgments of relative acceptability obtained for our Hebrew informants. Thus, even in languages in which objects of comparison can be relativized, there will be a preference to express the RC as one formed from a semantically equivalent sentence in which the semantic object of comparison is presented as a subject. That is, the preferred way to express the idea of the man who Mary is taller than will be, in general, the man who is shorter than Mary.

A somewhat more subtle performance preference is illustrated by the genitive position on the AH. Although a majority of languages in our sample possessed some way of relativizing genitives, there was often some awkwardness in doing so, and not infrequently, in specific cases, a preferred alternative was offered—one that relativizes
a position higher on the AH. For example, (58) is perfectly grammatical in French:

(58) la femme dont le manteau a été volé
    the woman whose the coat has been stolen
    'the woman whose coat was stolen'

But in practice people will tend to say:

(59) la femme qui s’est fait voler le manteau
    the woman who herself is caused to-steal the coat
    (59) is literally 'the woman who got her coat stolen', although there is no necessary implication of her having engineered the theft; relativization is on the subject, rather than the genitive. The translations above indicate that a similar alternative to genitive relativization exists in English. A like situation holds in Swedish, where the (b) alternative in each pair is preferred:

(60) a. kvinnan, vars kappa blev stulen
    woman whose coat was stolen
    b. kvinnan, som fick sin kappa stulen
    woman who got her coat stolen
(61) a. kvinnan, vars man är på sjukhuset
    woman whose husband is in hospital
    b. kvinnan, som har sin man på sjukhuset
    woman who has her husband in hospital

In Yoruba, inalienable possessives/genitives can often be paraphrased by other constructions, and even with simple sentences these paraphrases are preferred:

    John strike leg man the
    'John struck the man’s leg.’
    b. John lu ọkunrin naa l’ ṣe. ẹ
    John strike man the on leg
    'John struck the man on his leg.’

Under relativization, the preference is even more strongly marked, and our informant was very hesitant about admitting (63a) below:

    I see man that John strike leg his
    'I see the man whose leg John struck.’
    b. Mo ri ọkunrin ti John lu l’ ṣe.
    I see man that John strike on leg
    'I see the man that John struck on the leg.’
Bamgbose (1966, 158–159) claims that the relation between sentences like (63a) and (63b) is transformational.

2.2.2. Pronoun Retention. A less obvious patterning determined in part by the HCs concerns the distribution of personal pronouns in relativized positions. (We exclude pronouns that are instances of verb agreement; see section 1.4.2.2). We have already noted that Semitic languages “characteristically” present such pronouns, e.g. Hebrew:

(64) ha- isha she- David natan la et ha- sefer
    the woman that David gave to-her DO the book
    ‘the woman that David gave the book to’

Further, it has been argued in Keenan (1972a; 1975a) that such RC-forming strategies present in surface structure more of the logical structure of the RC than do languages like English that do not present such pronouns. The reason, in brief, is that in the pronoun-retaining strategies the restricting clause in surface is a sentence—one that expresses exactly the restricting sentence of logical structure. That is, it is just the sentence that must be true of the referent of the RC. So a possible referent of (64) above must be a woman of whom the sentence David gave the book to her is true. And the translation of that sentence in Hebrew is precisely David natan la et ha-sefer—the surface restricting clause. Note that the corresponding clause in English, David gave the book to or to whom David gave the book, is not a surface structure sentence, and so is not immediately perceived as the sort of linguistic entity that is true of objects.

It has been shown elsewhere (Keenan (1972a, and especially 1975a)) that the RC-forming strategies that retain pronouns are applicable to a greater range of otherwise “difficult” environments (e.g. it is often, but not always, possible in these languages to relativize into coordinate NPs, other relative clauses, indirect questions, and even sentence complements of NPs). The reason is that the logically more explicit strategies still successfully express the basic meaning of the RC in contexts where the meaning is otherwise difficult to perceive. Consequently, we are led to predict that, as we descend the AH, languages will exhibit a greater tendency to use pronoun-retaining RC-forming strategies. Table 2 (p. 93) dramatically supports this prediction. Languages that do not normally retain pronouns in any position are not included in the table; nor is the use of pronouns as markers of verb agreement (see 1.4.2.2).

It should be clear from Table 2 that not only does the tendency to present pronouns in positions relativized increase as we descend the AH, but also that once a language begins to retain pronouns it must do so for as long as relativization is possible at all. This is a natural consequence of the hypothesis that pronoun retention will be used in proportion to the difficulty of the position being relativized, though the critical point of difficulty is different for different languages.

2.2.3. Explaining the Psychological Validity of the AH. To some extent, explaining the HCs by interpreting the AH as a psychological hierarchy merely pushes back the
Table 2

Pattern of Pronoun Retention in Relative Clauses

<table>
<thead>
<tr>
<th>Language</th>
<th>Subj</th>
<th>DO</th>
<th>IO</th>
<th>Obl</th>
<th>Gen</th>
<th>OComp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aoban (North-East)</td>
<td>(+)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Arabic</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Batak</td>
<td>-</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Chinese (Pekingese)</td>
<td>-</td>
<td>+/−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Czech (colloquial)</td>
<td>-</td>
<td>+/−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Genoese</td>
<td>-</td>
<td>(+)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Gilbertese</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Greek (Modern)</td>
<td>-</td>
<td>-</td>
<td>+(?)</td>
<td>+(?)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Hausa</td>
<td>-</td>
<td>-</td>
<td>(+)</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hebrew</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Japanese</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>+/-</td>
<td></td>
</tr>
<tr>
<td>Javanese</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kera</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>*</td>
</tr>
<tr>
<td>Korean</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>Malay</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>*</td>
</tr>
<tr>
<td>Minang-Kabau</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>+/−</td>
<td>*</td>
</tr>
<tr>
<td>Persian</td>
<td>-</td>
<td>(+)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Roviana</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>Shona</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>(+)</td>
<td>+(?)</td>
</tr>
<tr>
<td>Slovenian</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Turkish</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Urhobo</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Welsh</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Yoruba</td>
<td>-</td>
<td>-</td>
<td>*</td>
<td>+</td>
<td>+</td>
<td>*</td>
</tr>
<tr>
<td>Zurich German</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Key: + means that personal pronouns are normally present in that position when it is relativized, using that RC-forming strategy which admits of pronoun retention. (+) means optional retention. +/- means that in some cases the pronoun is retained and in others it is not. - means that pronouns are usually not retained. * means that that NP position does not naturally exist in that language. 0 means that that position is not relativizable, and a blank means that we lack the relevant data. An entry of the form x(?) means that our data are uncertain but x is our best guess.

problem of explanation one step. We would still want to know why it is psychologically easier to relativize subjects than objects, etc. We have two speculations to make here.

2.2.3.1. A Recognition Strategy. Impressionistically, the initial portion of the AH appears to coincide with the degree to which NPs are required to appear in simple sentences. Thus, lexical predicates almost always require a subject. Many require
DOs, some require IOs, and a few (e.g. put) require OBLs. Furthermore, a few verbs (R. Stockwell, P. Schachter, personal communication) such as gnash (one's teeth), blink (one's eyes), and water (his eyes watered) appear to require that their arguments be possessed body parts, similar to the more idiomatic constructions lose one's nerve, blow one's cool, etc. Finally, no lexical predicate in English requires that it be construed in a comparative construction.

Perhaps, then, there is a kind of universal recognition strategy of the sort: "If an NP plays a role in another clause, interpret it as a subject unless there are indications to the contrary, otherwise try the DO slot, etc." Needless to say, such a general recognition strategy would require a more precise formation and much experimental research before it could constitute a serious explanation of the AH.

2.2.3.2. Independent Reference. Another explanation, which pertains only to the relative accessibility of subjects over other NPs, was offered in Keenan (1974). There it was argued that heads of RCs share a logical property with subjects of sentences but do not share this property with nonsubjects. Thus, more of what we need to know to understand the meaning of a RC formed on subjects is already contained in the meaning of simple sentences than is the case when the RC is formed on a nonsubject. In more "transformational" terms, RC formation on subjects distorts the meaning of the underlying sentence less than RCs formed on objects.

The logical property shared by heads and subjects is that of independent reference. Thus, in simple sentences we cannot generally make the reference of the subject NP dependent on that of some other NP in the sentence. For instance, if subject and object are marked as coreferential, it must be the object which is marked (if anything is), by for instance a reflexive pronoun. Thus, the reference of object phrases can be made dependent on that of subjects, but not conversely. But it is in the inherent nature of subjects to be independently referring. (For a weakening, but not abandonment, of this principle for more complex cases, see Keenan (1974).)

Similarly, to understand the meaning of an RC such as the girl that John likes, we must be able to understand what set is designated by the head NP independently of the reference of the NPs in the restricting clause. For instance, the head NP can never be coreferentially pronominalized by an NP in the restricting clause, even if, as in many languages, the restricting clause precedes the head. But of course NPs within the restricting clause can be stipulated as, for instance, being identical in reference to the head, as in (1), repeated below:

(65) der in seinem Büro arbeitende Mann
    the in his study working man
    'the man who is working in his study'

Consequently, if we relativize on a nonsubject, the resulting structure will contain two necessarily independently referring expressions, the head of the RC and the subject of the restricting clause. This explains, for instance, why the pronoun in the man who he
hit cannot be understood as being necessarily coreferential with the head. But if we relativize on a subject, there is only one necessarily independently referring expression. In this sense, then, subject relatives are psychologically simpler than nonsubject relatives.

3. Further Prospects

In this article, we have restricted ourselves by and large to the particular data area that originally led us to propose the Accessibility Hierarchy, namely, restrictions on relative clause formation. We wish here to indicate briefly two other areas where recent work has shown the possible relevance of the same Accessibility Hierarchy.

In Comrie (in press), it is shown that the AH is useful in the syntactic description of causative constructions, in particular synthetic causative formations, in a variety of languages. Summarizing the argument presented in the cited work, we may say that the syntactic position used to encode the causee of a causative construction (i.e. the individual caused to carry out some action) will be the highest position on the AH that is not already occupied. The following French examples illustrate the general trend of the data:

(66) J’ai fait courir Henriette.
    ‘I made Henriette run.’

In (66), the causee is presented as a DO (Henriette); note that the construction already has a subject.

(67) J’ai fait manger les gâteaux à Henriette.
    ‘I made Henriette eat the cakes.’

Here, the causee is presented as an IO (à Henriette); the construction already has a subject (je) and a DO (les gâteaux).

(68) J’ai fait écrire une lettre au directeur par Henriette.
    ‘I made Henriette write a letter to the director.’

In (68), the causee is an OBL (par Henriette); there is already a subject (je), a DO (une lettre), and an IO (au directeur).

A second extension of the range of applicability of the AH is in determining cross-language restrictions on advancement processes (Keenan and Comrie (1972); Perlmutter and Postal (1974); Johnson (1974b); Keenan (1975b); Trithart (1975)). By an “advancement process” we shall mean a productive syntactic process that converts sentences containing NPs on a low position on the AH into roughly synonymous

---

Similar facts obtain in languages in which the RC precedes the head. Thus in Japanese and Basque, for instance, in the equivalent of “(the) he hit man” (= “the man that he hit”), the subject pronoun of the main verb in the RC cannot be understood as being coreferential with the head. See Keenan (1974) for examples. Thus the Crossover Principle (Postal (1971)) is not the correct explanation for these cases.
sentences containing that NP on a higher position on the AH. An obvious example would be Passive, which in many languages serves to advance direct objects to subject position. The work cited above leads us to posit two Advancement Conditions:

(AC1) The Target Condition
The higher an NP is on the AH, the more accessible it is, in general, as a target of an advancement.

(AC2) The Distance Condition
The farther X is from a fixed position Y higher than X, the more difficult it is, in general, to advance X to Y.

(The qualification "in general" is required because many such advancement processes, though productive, are subject to idiosyncratic government by individual lexical items.) The Target Condition would claim, for instance, that if a language has a rule advancing non-DOs to DO, then it must also have a rule advancing non-Subjects to Subject; that is, since DO is a target for such processes, there must also be processes with Subject as their target. Indonesian (Chung (1976)) is an instance of a language with advancement of IO to DO, and, as predicted by AC₁, it also has advancement of DO to Subject. The Distance Condition would claim, for instance, that if a language has a rule advancing OBL to Subject, then it must also have a rule advancing DO to SU; that is, since OBL is farther away from Subject than DO is, OBL is a less likely candidate for advancement than DO. Malagasy (Keenan (1972b)) has a rule advancing OBLs (and also IOs) to Subject, and also has a rule, as predicted, advancing DO to Subject. For further examples and detailed justification, the works cited above should be referred to.

In conclusion, this article has aimed to demonstrate that the Accessibility Hierarchy correctly characterizes the relative accessibility of different syntactic positions to relative clause formation; other work indicates that the Accessibility Hierarchy may play a more general role in determining the accessibility of noun phrases as candidates and targets for syntactic processes, although the precise delimitation of the area of relevance of the Accessibility Hierarchy remains a task for future research.

References
Cook, V. J. (forthcoming) "Strategies in the Comprehension of Relative Clauses," *Language and Speech*.
Gorbet, L. (1972) "How to Tell a Head When You See One: Disambiguation in Digueño Relative Clauses," unpublished manuscript.


Legum, S. (1975) "Strategies in the Acquisition of Relative Clauses," Technical Note 2-75-10, Southwest Regional Laboratory, Los Alamitos, California.


Valli, H., N. Hernandez, M. Archard-Baule, and M. F. Beretti (1972) Compte-rendu d'une expérience réalisée dans une classe de 6e 2, dont le but était d'étudier les mécanismes de la production des relatives chez l'enfant, Institut de Didactique et Pédagogie, Université de Provence.

Keenan
Department of Linguistics
UCLA
Los Angeles, California 90024

Comrie
Department of Linguistics
Sidgwick Avenue
Cambridge CB3 9DA
England