2) it specifies the object of investigation for the particular linguistic subdisciplines from one superordinate linguistic vantage point;
3) it does not introduce a hierarchy of description for the theory of language, it points to a dialectical interrelation between all components of language and all areas of language research;
4) it equally sanctions both methods of scientific access to objective truth, that is the method of induction (experimenting) and deduction (logic);
5) it confirms the humanistic character of linguistics, which describes both that which is predictable (which coincides more or less with the social) and that which is unpredictable (which coincides more or less with the individual);
6) it tolerates all linguistic trends. It evaluates them only with respect to verifiability of theoretical rules with linguistic reality.

Thus linguistics should study the laws governing the linguistic reality in its total design, and overall structure. Only in this sense can linguistics become a universal science giving hope for a better and better insight into the phenomenon which is reputed to be the most unique — the human language.

Definite and Animate Direct Objects: A Natural Class*

BERNARD COMRIE
King's College, Cambridge, England

0. The data base of the present paper comes primarily from case--marking of direct objects in a variety of languages. In a number of languages, one finds a special marker for direct objects which are definite[1]; other direct objects have the same form as the nominative, typically with no marker at all. Examples (1) and (2) below are from Persian and Turkish, respectively; similar examples can be drawn from most Turkic languages, and also from Uralic languages (Wickman 1955; Comrie, forthcoming):

(1) a. Ketâb-râ didam.
book DO I-saw
'I saw the book.'
b. Ketâb-i didam.
book a I-saw
'I saw a book.'

(Note that -i is a marker of the indefinite article irrespective of syntactic position, i.e. does not mark specifically indefinite objects.)

* This work was supported in part by a grant from the Social Science Research Council for the investigation of the syntactic typology of the non-Slavic languages of the Soviet Union.

[1] In some languages, the relevant characteristic is not so much definiteness, in the strict sense, as specificity, and a number of related concepts. For an argument that all of these can be related, see Comrie (in preparation). For present purposes, we shall use the term 'definite' to cover all of these. In some languages, topicality (thematicity) is also relevant.

(2)

a. Hasan öküz-ü gördü.
Hasan ox Do saw
'Hasan saw the ox.'

b. Hasan bir öküz gördü.
Hasan a ox saw
'Hasan saw an ox.'

In many languages, one finds a special marker for direct objects that are animate (or, more restrictively, that are human), for instance in Russian[2],

(3)

a. Begemont ljubit nosorog -a.
hippopotamus loves rhinoceros DO
'The/a hippopotamus loves the/a rhinoceros.'

b. Begemont ljubit il.
hippopotamus loves slime
'The hippopotamus loves (the) slime.'

1.1 So far, we have simply noted that some languages have the phenomenon of definite direct object marking, and some have the phenomenon of animate direct object marking. However, the interesting feature is that in many, if not most, of these languages there is no absolutely rigid distinction between definite direct object marking and animate direct object marking. In some languages the 'seepage' between the two phenomena is slight. In Persian and Turkish (and other Turkic languages),for instance, the animate interrogative pronouns *ki* and *kim*, respectively, 'who?' require the definite direct object ending when occurring as direct object:

(4)

Ki- râ didam?
who DO I-saw
'Who did I see?'

(5)

Hasan kim-i gördü?
Hasan who DO saw
'Who did Hasan see?'

Thus *-râ* and *-i/-i/-u/-ü* (vowel harmony variants), which usually

[2] More accurately: In Russian, animate direct objects take a special marker, that of the genitive case, unless they belong to a declensional type that has a separate accusative case distinct from all other case-forms; inanimate nouns lacking a special accusative have the accusative like the nominative; i.e. the special marker is used only where the noun phrase is animate and where otherwise it would have the same form as the nominative. For an explanation for this, see section 2 below.

mark definite direct objects, here mark animate direct objects[3]. In Russian, the genitive-like accusative usually marks animate direct objects, but there is one instance where it is used to mark inanimate definite direct objects, namely with the pronouns 'it', 'them', which have nominative *on* (masculine), *ona* (feminine), *ono* (neuter), *oni* (plural), but accusative = genitive *ego* (masculine and neuter), *ee* (feminine), *ix* (plural):

(6)

Begemont ljubit ego (= nosorog-a, il).
'The hippopotamus loves it (=the rhinoceros, the slime).'

1.2 In many languages, the seepage between definite direct object marking and animate object marking is even greater. In one type, exhibited for instance by Spanish, both definiteness (in Spanish: specificity) and animacy are required for the special direct object marker (preposition *a*) to be used:

(7)

a. El director busca un automóvil.
'The director is looking for a car.'

b. El director busca el automóvil.
'The director is looking for the car.'

c. El director busca un empleado.
'The director is looking for a clerk (i.e. any-clerk).'

d. El director busca al (= a el) empleado.
'The director is looking for the clerk.'

In the closely related Catalan language, the obligatory use of *a* is much more restricted (although extensions of the obligatory use are often found, probably under the influence of Spanish): it is found only with first and second person pronouns. Note that such pronouns are necessarily animate and definite, and as such can be seen to represent the core of the intersection of animacy and definiteness:

(8)

a. No m'havien vist a mi.
'They had not seen me.'

b. No havien vist l'alcalde.
'They had not seen the mayor.'

Similarly in the completely unrelated language Finnish, the accusative in *-t* is used only for personal pronouns referring to humans (i.e.

[3] Alternatively, one could argue that the human pronoun is necessarily definite, or rather specific, as is done by Browne (1970 : 362). Either way, a relationship is established between 'human' and 'specific', i.e. between 'animate' and 'definite'.

including *hän*, accusative *hänet* 'he, she,' but not *se* 'it'), and also for the human interrogative pronoun *kuka*, accusative *kenet* 'who?':

(9)
a. Minä näin hänet.
'I saw him.'
b. Kenet minä näin?
'Who did I see?'

In the Median dialects spoken around Qazvin in northwestern Iran and described by Yar-Shater (1969), who refers to them as Southern Tati, we find that there is a special marker in the singular for definite direct objects. Although this marker is used only with definite direct objects, it is not in fact used with all definite direct objects. It is, however, used with all definite human direct objects, and apparently nearly always with other definite animate direct objects; with inanimate definite direct objects and occasionally with nonhuman animates, the absolute case is used when the direct object is qualified by a genitive or possessive adjective, and the oblique case otherwise, e.g. (Yar-Shater 1969: 98—100) in the Xiāraji dialect (oblique ending *-e*):

(10)
a. Pëvl Hasan-e di.
money Hasan to give
'Give money to Hasan.'
b. Pëvl-e Hasan-e di.
'Give the money to Hasan.'
c. Cemā qoc-e hunzan.
'Call our ram.'
d. Cemen jurab (ABS) pináka.
'Patch my sock.'

1.3. Indeed, when we turn to descriptive grammars of some languages, it is very difficult to decide whether we are dealing with definite direct object marking, or animate direct object marking: often different grammars of the same language assign priority to a different feature of direct objects, often the same descriptive grammar will present an account which, if taken literally, is confused or self-contradictory. Below, I give examples from three languages (Hindi, Tagalog, Mongolian) to demonstrate the intersection of animacy and definiteness in direct object marking; obviously I have chosen the data sources cited below because they are reasonably clear expositions, even if not always complete, so that the structures voiced above against certain descriptive grammars do not apply to those cited below.

In Hindi, the postposition *ko* marks definite/animate (especially human) direct objects. Looking at the examples in McGregor (1972 : 48), however, it is clear that there is no simple correlation of *ko* with either animate or definite direct objects:

(11)
a. Aurat bacce ko bulā rahi hai.
woman child DO calling PROG is
'The woman is calling the/a child.'
b. ?Aurat baccā bulā rahī hai.
'The woman is calling a child.'
c. Darzī bulāo.
tailor call
'Call a tailor.'
d. Un patrōm ko parhie.
those letters DO read (POL)
'Please read those letters.'
e) Ye patr parhie.
these letters read (POL)
'Please read these letters.'
f) Patr likhie.
letter write (POL)
'Please write a letter.'

(Note that some nouns have distinct forms before postpositions, e.g. *baccā* 'child' is *bacce* before a postposition.) If we take animacy as the decisive criterion, then we have to admit that inanimate nouns, if definite, also allow *ko*, and that indefinite animate nouns, especially if nonspecific, allow omission of *ko*. If we take definiteness as the decisive criterion, then we have to allow that indefinite, but specific, animate direct objects can take *ko*, whereas definite inanimate direct objects can occur without *ko*. In fact, both definiteness and animacy are relevant.

In Tagalog, the preposition *ng* (before personal proper names: *ni*) is used before underlying subjects (agents) and direct objects that are not topicalized; the preposition *sa* (before personal proper names: *kay*) before indirect objects; and the preposition *ang* (before personal proper names: *si*) before topics. In general, definite direct objects must be topicalized in Tagalog, so that the question of a marker for definite direct objects does not arise. However, there is one construction where this topicalization is not possible, the 'nominalized verbal' (Schachter and Otanes 1972: 382—383, with the meaning 'he/she is the one who...', and here both *ng* and *sa/kay* are found marking the direct object. If the direct object is indefinite, only *ng* is possible. The preposition *sa/kay* is thus possible only with definite direct objects, and is in fact obligatory with personal proper names and pronouns (the class of noun phrases which are necessarily definite and animate), possible with other

human noun phrases, and possible, though rare, with inanimate direct objects; where *sa/kay* is not used, *ng* is used:

(12)
a. Siya ang nakakita kay Jose.
he TOP the-one-who-saw Jose
'He is the one who saw Jose.'
b. Siya ang nakakita sa/ng duktor.
'He is the one who saw the doctor.'
c. Siya ang nakakita ng duktor.
'He is the one who saw the doctor.'
d. Siya ang nakakita ng/? sa aksidente.[4]
'He is the one who saw the doctor.'
e. Siya ang nakakita ng aksidente.
'He is the one who saw the accident.'

Again, both animacy and definiteness are relevant, though possibly with some preponderance of definiteness (indefinite direct objects always have *ng*).

In Mongolian,[5] the accusative suffix *-(ii)g* is added to certain direct objects only (Poppe 1970: 148—149), though it is more frequent than, say, in Turkish. Where the direct object is human, the special accusative case must be used:

(13)
a. Dorž bagš -iig zalav.
Dorj teacher DO invited
'Dorj invited the teacher.'
b. Bid nar olan xün -iig üzsen.
we many people DO saw
'We saw many people.'

Where the direct object is nonhuman, both accusative and nonaccusative are found, with the definiteness of the direct object (and also its thematicity) being an important factor in conditioning the accusative, although the accusative is also found with indefinite direct objects, particularly if they are separated from the verb:

[4] My informant rejected the variant of (12d) with *sa*, while Schachter and Otanes (1972 : 383) describe it as less preferred. In Tagalog, and persumably in the other languages cited, there is thus some idiolectal variation in the precise weighting of definiteness and animacy in thus determining the use of the special direct object form. Thus Rekha Saith informs me that Hindi sentence (11b) is very marginal.

[5] The examples given here are all from Khalkha-Mongolian, the standard language of Mongolia proper, but the use of the accusative is essentially the same in other Mongolian languages/dialects.

(14)
a. Čoidog zurag zurav.
Choidog picture painted
'Choidog painted a picture.'
b. Zurag -iig Čoidog zurav.
picture DO Choidog painted
'Choidog painted the picture; as for the picture, it was Choidog that painted it.'

Thus again, both animacy and definiteness are relevant criteria in the choice of the special direct object suffix.

2. The data presented above, from a variety of languages, serve to show that definite and animate direct objects often function together for the purposes of case-marking, i.e. form a natural class (or at least: form a natural class for the purpose of the rule of case-marking). We may now turn to trying to explain the basis of this natural class, i.e. to trying to find what feature animate and definite direct objects have in common that leads to their being treated similarly by case-marking. As the basis of our explanation we shall take the observation that, in natural languages, certain grammatical relations tend to be characterized by certain features, in particular: that subjects tend to be definite, animate, and topic (thematic); while direct objects tend to be indefinite, inanimate, and rhematic (Givón, forthcoming; Keenan, forthcoming).[6] Secondly, we shall assume that the essential function of case-marking of subjects and direct objects is not so much to have an overt marking of subjects and to have an overt marking of direct objects, but rather to have an overt marking of the *difference* between subjects and direct objects; for further development of this approach, using a more restricted data base, see Comrie (forthcoming). From the second of these postulates, it would follow that overt case-marking of a direct object with a form distinct from that of a subject would be particularly likely where there is a greater likelihood of confusion between subject and direct object. Since subjects are typically animate and definite, one kind of direct object that is particularly likely to get a special marker will be animate and/or definite direct objects. Thus the feature that animate and definite direct objects have in common, which leads to their being marked in the same way in so many languages, is that both are kinds of direct object that are particularly likely to be confused with subjects unless they are so marked.

3. Finally, we may look at the relationship between our account of case-marking of definite/animate direct objects and recent work on

[6] We shall not consider further here the possibility that 'subject' might be *defined* in terms of a weighting of its characteristic features, though a promising approach along these lines is developed by Keenan (forthcoming). For our purposes, it is necessary only to assume that subjects typically have these properties, and similarly for direct objects.

verb-object agreement, in particular Givón (forthcoming). Givón notes that, in natural languages, the arguments with which verbs tend to show argument are (i) subjects, (ii) definite and/or animate direct objects.[7] Thus we can say, more generally, that verbs tend to agree with those arguments that are, or at least: are typically, definite, animate, and thematic. One might therefore be tempted to subsume both case-marking and verb-agreement under the same unified general explanation. We suggest that this cannot be done, although the fact that subjects tend to be definite, animate, and thematic is no doubt an integral part of the explanation for both phenomena. The crucial difference, however, lies in the *discriminatory* role of case-marking of definite/animate directs objects: they are marked differently from subjects, although they share certain semantic features in common, indeed *because* they share certain features in common, as the function of this case-marking is to keep them apart overtly. Verb-agreement, however, has little or no discriminatory function, indeed it serves to *unite* definite/animate direct objects and subjects.[8] A further argument concerns case-marking of subjects. Although in this paper we are concerned primarily with direct objects, there is some evidence that languages also tend sometimes to have special markers for subjects which are particularly likely to be confused with direct objects (e.g. where the simplex sentence also contains a direct object, or where the subject is indefinite) (Comrie 1975; forthcoming). However, we are not aware of any language where subject-verb agreement is used to mark indefinite subjects, although this would be predicted if verb-agreement and case-marking functioned exactly in parallel;[9] indeed, in many languages verb-subject agreement is, or can be, suspended precisely where the subject is indefinite, e.g. French:

(15)
a. Les arbres sont (PL) dans le jardin.
'The trees are in the garden.'
b. Il est (SG)/il y a (SG) des arbres dans le jardin.
'There are (colloquially: theres's)some trees in the garden.'

[7] Actually, more generally definite/animate nonoblique objects, since indirect objects and benefactives are often, though not invariably, included. Note that the special case-marking for definite/animate direct objects is often also (at least etymologically) an indirect object marker, e.g. Persian *-râ*, Spanish *a*, Hindi *ko*, Tagalog *sa/kay*. Indirect objects and benefactives are, of course, typically animate; see Givón (forthcoming).

[8] In general, verb-agreement markers of subject-agreement and object-agreement are distinct, though not invariably so, in particular in languages (such as many ergative languages) where verbs agree with intransitive subjects and (some) transitive direct objects, but not with transitive subjects.

[9] I am grateful to Katalin Radics for pointing this consequence out to me.

4. Conclusion

In addition to its parochial relevance to the problem of case-marking of direct objects, the line of analysis pursued in this paper has some more basic implications for linguistic theory. Firstly, it emphasizes the role played by morphology in discriminating between different categories, rather than in simply encoding categories. Secondly, it suggests that in constructing explanatory hypotheses one must ask not only whether a given element is characterized by a certain feature (e.g. whether the direct object of a sentence is definite or not), but also whether elements of that class are typically, independent of the given sentence, characterized by a certain feature (e.g. whether direct objects are typically definite or not).

References

Browne, E. W. 1970. *More on Definiteness Markers: Interrogatives in Persian,* "Linguistic Inquiry," 1, pp. 359—363.

Comrie, B. 1975. *The Antiergative: Finland's Answer to Basque,* "Chicago Linguistic Society," 11, pp. 112—121.

Comrie B. Forthcoming. *Subjects and Direct Objects in Uralic Languages: A. Functional Explanation of Case-Marking Systems,* "Etudes Finno-Ougriennes," 12.

Comrie, B. In preparation. *'Definite' Direct Objects and Referent Identifiability.*

Givón, T. Forthcoming. *Topic, Pronoun and Grammatical Agreement,* [in:] C. Li, e.d., *Subject and Topic.*

Keenan, E. L. Forthcoming. *Towards a Universal Definition of 'Subject of',* [in:], C. Li, ed., *Subject and Topic.*

McGregor, R. S. 1972. *Outline of Hindi Grammar,* Oxford: Clarendon Press.

Poppe, N. 1970. *Mongolian Language Handbook,* Washington D.C.: Center for Applied Linguistics.

Schachter, P. and F. T. Otanes. 1972. *Tagalog Reference Grammar,* Berkeley: University of California Press.

Wickman, B. 1955. *The Form of the Object in the Uralic Languages,* Uppsala: Almqvist and Wiksell.

Yar-Shater, E. 1969. *A Grammar of Southern Tati Dialects,* The Hague: Mouton.