Agreement in Natural Language

APPROACHES, THEORIES, DESCRIPTIONS

EDITED BY MICHAEL BARLOW & CHARLES A. FERGUSON

CSLI CENTER FOR THE STUDY OF LANGUAGE AND INFORMATION

1988

7 Agreement vs. Case Marking and Direct Objects

WILLIAM CROFT

THIS PAPER BEGINS with a problem in the characterization of the behavior of direct objects with respect to agreement and case marking across natural languages, and in the process of developing a solution to this problem I propose an analysis of the presence vs. the absence of agreement and case marking of arguments. The form of this analysis is that of universal structure–meaning (form–function) correspondences for agreement and case marking. This is a new type of constraint on possible human language types, and it implies that the cognitive processes which must account for such constraints are quite general. The analysis also unifies the three most important typological patterns that link structure and function in natural languages—markedness, hierarchies and prototypes—under a single concept, that of *relative markedness*, thereby allowing typological theory to be characterized in a more general and perspicuous manner.

I would like to thank Joseph H. Greenberg, Charles Ferguson and Elizabeth Traugott for their comments on an earlier version of this paper. I would also like to thank Carolyn Coleman for providing the Kun-parlang data, and Tony Davis for discussion of the Hausa data. The Quiché data is from materials in a course on Quiché offered by Norman McQuown at the University of Chicago in 1976-77.

The problem, which has been prominent in recent literature on grammatical relations and transitivity, is: are definite and/or animate direct objects marked or unmarked?

Hopper and Thompson (1980) and Givón (1976) suggest that they are unmarked. Hopper and Thompson propose an abstract concept of Transitivity to which a large number of properties are correlated. If one assumes that the "ideal" transitive utterance is the unmarked one, as Hopper and Thompson seem to imply, then one would predict that the animate and definite direct object will be unmarked, since it is highly individuated (Hopper and Thompson 1980:252-253). Givón is more explicit: he describes agreement as obeying a hierarchy Agent < Dative < Patient, the verb agreeing with the arguments that are higher on the hierarchy (Givón 1976:152, 160-166). Givón's hierarchy inverts two members of the standard case hierarchy (Keenan and Comrie 1977), in which the accusative precedes the dative. If Givón is right, then the concept of a universal case hierarchy is badly damaged.

In contrast to the position of Hopper and Thompson and Givón, Comrie (1979) suggests that definite/animate direct objects are marked, chiefly on the basis of the presence vs. absence of case marking on direct objects in a variety of languages. Thus, we have two conflicting hypotheses. The evidence that Comrie, Givón, and Hopper and Thompson cite is all descriptively accurate, however. It appears that in this case, marking theory provides inconclusive or conflicting evidence concerning the status of certain kinds of direct objects. This means that either there is no universal characterization of the markedness of direct objects, or that there is some deficiency in marking theory. I will argue that the problem can be solved by successive extensions to classical marking theory, as originally developed by Trubetzkoy, Jakobson and Greenberg. I will also argue that ultimately Comrie is correct, but the phenomena that Hopper and Thompson and Givón cite can also be accounted for.

1 The Markedness of Direct Objects

1.1 Classical Marking Theory

The classical theory of markedness is based on the discovery that paradigmatic members of the same grammatical category have asymmetrical linguistic properties. The classical theory of markedness is based largely on three properties: (1) unmarked values are morphosyntactically less complex than marked ones, in fact, they are usually "zero-marked"; (2) marked values are behaviorally defective compared to unmarked ones, that is, they do not inflect for as many grammatical categories (e.g., number) and/or their syntactic distribution is more restricted; and (3) unmarked values are textually more frequent than marked ones. In this

paper, I will present arguments based chiefly on the first criterion and partly on the second; see Greenberg 1966 for extensive studies based on the third.[1]

Formally, we can represent the paradigmatic elements as privileged members of a single set, the set of members of a grammatical category:

$$C = \{u, m\} \qquad \text{OR} \qquad C = \{u, m_1, m_2, \ldots\}$$

($u, m_1, m_2, \ldots$ elements of a category C)

In the classical theory of markedness, the solution to the problem we have posed seems quite simple: definite and/or animate direct objects are marked, because if there is a contrast between the presence of an agreement marker or a case marker and its absence in the direct object, it is always the more definite and/or more animate direct object which is marked by agreement and case. This is true across a very large number of languages and has been widely observed in the literature. We will consider some typical cross-linguistic facts.

In Swahili (Perrott 1972:38), agreement is sensitive to definiteness. The verb agrees with definite direct objects but not indefinite ones (there is evidence of interactions with animacy as well; see Givón 1976:159, Ashton 1944:58, 60):

(1) U- me- leta kitabu?
2.SG- PERF- bring book
'Have you brought a book?'

(2) U- me- ki- leta kitabu?
2.SG- PERF- 3.SG- bring book
'Have you brought the book?'

Turning to case marking, Punjabi (Shackle 1972:69-70) provides an example of a language in which both animacy and definiteness interact with direct object status. First and second person direct objects, which are always animate and always definite by definition, require the dative postposition; third person direct objects require the postposition if they are definite, and do not use it when they are indefinite:

(3) mɛ̃ tɛ -nũ pəṛàvaŋga
1.SG 2.SG -DAT will.teach
'I will teach you.'

[1] The version of the theory of markedness used in generative grammatical theory, particularly in generative phonology, is a different mechanism. The generative phonological use of markedness is related to the first criterion, namely, to represent "default" feature values to simplify abstract (phonological) representations and rules.

(4) ó nili kitāb nū mez te rakkho
that blue book to table on put
'Put that blue book on the table.'

(5) koi kitāb mez te rakkho
some book table on put
'Put some book on the table.'

Comrie, in an extended discussion of the interaction of animacy and definiteness with case marking (Comrie 1979), also cites Persian, Hindi, Turkish, Russian, Spanish, Tagalog and Mongolian as languages in which both animacy and definiteness govern variation in case marking.

1.2 Marking Theory and Grammatical Hierarchies

While classical marking theory can account for the widespread data of the kind presented in the Swahili and Punjabi examples, there are other grammatical phenomena which are obviously quite closely related, but classical marking theory cannot handle them without modifications.

Kun-parlang, an Australian language, agreement with the direct object interacts with animacy. The probability of the verb agreeing with its direct object depends on its animacy: the verb agrees with a first or second person direct object almost always, with a third person human object often, and with a third person inanimate direct object almost never:

(6) nga- ngum- kinyang
1.SG.REAL- 2.SG- cook.PAST
'I burned you.'

(7) nga- kinyang
1.SG.REAL- cook.PAST
'I burned it/him'

Thus, there is a gradation of "degree" of animacy that determines agreement in Kun-parlang.

Another example of gradable determination, this time with case marking, is found in Rumanian (Nandris 1945:183-185). Rumanian is still in the process of losing its last case distinction inherited from Latin (nominative vs. oblique), but has begun using prepositions for many of the grammatical functions, including using *pe* 'on' for direct objects. While the use of *pe* for direct objects is not easily described, there are certain situations in which *pe* is required, situations in which it is optional, and situations in which it is prohibited. It is required when the direct object is both human and definite (including pronouns), and with certain definite constructions (e.g., demonstrative + cardinal numeral, and ordinals); it is optional when the direct object is either human but a specific indefinite, or nonhuman but pronominal, and it is prohibited when the

direct object is a nonspecific indefinite or a generic. Thus, one must use *pe* when the direct object is both human and definite, one may use it when the direct object is either human or definite but not both, and one must not use it when it is (nonspecific) indefinite or generic.

The theory of markedness can be extended to handle more data such as the Kun-parlang and Rumanian examples by relativizing the notion of markedness. The classical theory of markedness states that a member of a category is either marked or unmarked, absolutely. But in certain cases, there is a scale on which members of the category can be placed, and the scale applies to linguistic phenomena which are clearly in the domain of marking theory, such as the presence vs. absence of agreement or case marking. The solution is to redefine the markedness of a member of a category *relative to* other members of the same category. Formally, this is represented by a (partial) ordering of the elements of the category:

$$C = \{m_1 < m_2, m_3 < m_4 < \ldots\}$$

The formal description given here is in the familiar form of an implicational hierarchy. Given some grammatical phenomenon, one can convert the $<$'s to $\subset$'s and read the implications from right to left to obtain a series of implicational universals. This relationship between markedness and implicational universals is not accidental, and has been discussed by Greenberg (1966:21-22). Greenberg also appears to be the first to have observed that implicational hierarchies are related to properties of markedness, in particular, the behavioral and textual (frequency) criteria (Greenberg 1966:31-32, 42-45). In this section, we will demonstrate that hierarchies also adhere to the structural criteria of markedness.

All of the categories we are concerned with—definiteness, animacy and case or grammatical relations[2]—have been found to form hierarchies with respect to various grammatical processes. They are: the Animacy Hierarchy (Silverstein 1976; Dixon 1979),[3] the Definiteness Hierarchy (Greenberg 1978), and the Case or Accessibility Hierarchy (Keenan and Comrie 1977), usually given as follows:

Animacy: 1st, 2nd < 3rd < Proper Name < Human < Animate < Inanimate

[2] The relationship between case and grammatical relations is quite controversial; I take the position that abstracting two (or more) separate levels is unnecessary (see Croft 1983). It is also true that most advocates of the "case" hierarchy have postulated only one hierarchy.

[3] The "Animacy Hierarchy" involves not only the animacy of the referent but also the speech act status (1st/2nd person versus 3rd person) and the type of referring expression (pronoun, proper name, or common noun). These factors are interdependent, however: 1st and 2nd person are always human and pronominal, and proper names are usually human. For this reason, the Animacy Hierarchy is normally treated as a unified phenomenon, and I will follow that practice here.

Definiteness: Definite < Specific/Referential Indefinite < Non-specific/Generic

Case: Subject < Direct Object < Indirect Object < Oblique

The extension of marking theory to hierarchies allows us to account for considerably more data. The presence vs. absence of agreement and case marking is to a great extent a function of the case hierarchy. Moravcsik (1974) discovered that agreement is associated with the upper end of the case hierarchy in the following way: if there is a construction in which the verb agrees with some member of the case hierarchy, then there are at least some constructions in which the verb agrees with members higher on the case hierarchy. This statement allows for the large number of languages which have variable object agreement, i.e., under certain conditions the verb agrees with indirect or oblique objects but not direct objects in a given construction, although it agrees with direct objects in other constructions. These languages themselves make up a subset which obeys the case hierarchy. An enumeration of existing types of agreement systems is given below:

1. Languages having no agreement: Lahu (Matisoff 1973), Chrau (Thomas 1971), Mandarin Chinese
2. Languages having agreement with the subject only: English, Russian, Turkish (Lewis 1967)
3. Languages having agreement with the subject and one object:
   a. patient only: Quiché, Ayacucho Quechua (Parker 1969)
   b. patient or Dat/Ben/Mal: Classical Nahuatl (A. Anderson 1973)
   c. patient, Dat/Ben/Mal, or oblique(s): Acoma (patient possessor; Miller 1965), Kun-parlang (comitative), Amharic (instrument; Moravcsik 1974:40)
4. Languages having agreement with subject, direct object, and indirect object or other oblique: Abkhaz (Hewitt 1979), Manam (Lichtenberk 1983)
5. Languages having agreement with subject, direct object, indirect object and benefactive: Kinyarwanda (Dryer 1983)

The presence of case marking, on the other hand, is associated with the lower end of the case hierarchy; that is to say, the absence of case marking is associated with the upper end of the hierarchy. Unlike agreement, there is a single simple hierarchy:

1. Languages with no zero case marking: Latvian (Ladziņa 1966), Japanese (Clark and Hamamura 1981)

2. Languages with zero subject (nominative/absolutive)[4] case marking: Hungarian (Whitney 1944), Turkish (Lewis 1967), Dyirbal (Dixon 1972)
3. Languages with zero subject and object case marking: Quiché, Persian (Mace 1962), Chrau (Thomas 1971)
4. Languages with zero subject, object and indirect object case marking: Manam (Lichtenberk 1983a), English

There are other phenomena outside the realm of agreement and case marking which supports the existence of the case hierarchy, namely relativization (Keenan and Comrie 1977) and causativization (Comrie 1976). Linking marking theory to these phenomena considerably broadens its scope of applicability.

Although the typological facts clearly require this extension of marking theory, we run into severe problems when we apply hierarchies to the phenomenon of definite/animate direct objects. Consider the introduction of animacy and definiteness as a partition of the case hierarchy, so that what we seek is an ordering within the category of "direct object" of definite/indefinite and animate/inanimate direct objects. Since agreement is associated with the upper end of the case hierarchy, presence of agreement implies that the entity is "less" marked. Hence, given the data above, definite/animate direct objects are less marked than indefinite/inanimate ones. Case marking, on the other hand, is associated with the lower end of the case hierarchy; therefore, presence of case marking implies that the entity is "more" marked. Thus, definite/animate direct objects are more marked than indefinite/inanimate ones. An anomaly not present in the classical theory has been generated by this extension of marking theory. Fortunately, there is a further extension to marking theory which allows us to resolve this anomaly and cover a still greater number of cross-linguistic phenomena.

1.3 Marking Theory and Natural Correlations

Actually, the original statement of the problem—are definite/animate direct objects marked or unmarked?—is an incoherent question in terms

[4] In ergative languages, the unmarked case in transitive constructions marks the direct object and in intransitive constructions, the subject. This would suggest a partial ordering Intr. Subject < Tr. Subject, Tr. Object < . . . . However, all languages possessing ergative patterns display "split ergativity" (Dixon 1979), in which the transitive subject is unmarked under certain conditions and even the intransitive subject is marked under certain conditions (so-called "active" languages). This anomaly for the case hierarchy is solved by the extension of marking theory to natural correlations in Section 2.3.

of both classical marking theory and the extension to grammatical hierarchies. Markedness is defined solely in terms of relationships among members of the *same* grammatical category. However, the problem under consideration mixes three *different* categories: case (associated with whatever relational morphosyntax is used), definiteness (associated with the determiner system), and animacy (associated with nouns themselves and the gender/class system, if there is one). One must extend the concept of markedness still further, to include cross-categorial relations.

The final definition of markedness includes the concept of an element of a grammatical category being unmarked *relative to a member of another grammatical category.* This unmarked correlation is called a *natural correlation.* Formally, this can be represented as sets of ordered n-tuples of correlated members of different categories:

Category	Values	
A	a	b
B	j	k
C	x	y

$< a, j, x >$: unmarked (natural) correlation

$< a, k, y >$: marked correlation

The cases of simple markedness and hierarchies discussed in the previous sections occur when there is no such markedness relationship across categories.[5] In a given utterance, a category value such as "direct object (patient)" is unmarked in its realization only if all of its natural correlations are present also; in other words, one can be certain that the construction "normally" or "typically" associated with the patient will actually be used to indicate the patient only when the properties which naturally correlate with patienthood are also present. When one or all of the cooccurring properties are not naturally correlated with the patient, then the linguistic manifestation of the patient is more marked, and thus the patient nominal may carry a grammatical mark, e.g., a (nonaccusative) case marking.

Natural correlations themselves must be explained in terms of properties external to the structure of language, e.g., certain typical correlations of the kinds of phenomena people normally talk about, and the amount of attention directed to different aspects of the phenomena

[5] Again, Greenberg appears to be the first to have suggested the possibility of markedness relative to values across grammatical categories: "It should be noted that in some cases we had what might be called conditional categories for marked and unmarked. For example, whereas for obstruents, voicing seems clearly the marked characteristic, for sonants the unvoiced feature has many of the qualities of a marked category" (Greenberg 1966:24).

(for example, human actors). The same is of course true of classical markedness and hierarchies within morphosyntactic categories. The relevance of natural correlations to the structure of language, and thus to grammatical theory, is manifold. They allow one to determine the "true" meaning of surface grammatical categories like "direct object" or "agreement," by specifying the conditions under which those meanings are expected to appear (that is, when their natural correlates are also present). This in turn allows us to define what aspects of the grammar will be motivated by (external) semantic and pragmatic factors and what aspects will be partially arbitrary conventionalizations in the grammars of specific languages.

The natural correlations themselves are hypothesized to be universal, and thereby externally motivated. When the meaning of a grammatical category is conventionalized (grammaticalized) in some natural language, e.g., surface direct objects are always indefinite, the grammaticalization is predicted to always align itself with its natural correlations. The arbitrary aspect of grammar is the degree to which a language will mark "unnatural" correlations. Languages may or may not mark "unnatural" correlations with a distinct surface form, and for that reason, cross-linguistic variation in the marking of, say, direct objects (patients) is found. Even then, the implicational scales determined by the hierarchies are adhered to within individual languages. In these respects, natural correlations behave like "core uses," "prototypes," or "ideal types." The link between marking theory and prototypes allows us to integrate prototype analyses of certain grammatical categories with more traditional morphosyntactic properties of those categories.

It is clear from the evidence presented so far that agreement markers and case markers have a different set of natural correlations, that is, agreement and case marking have universal but distinct properties. So, it is possible to assign a "meaning" or significance to the presence vs. absence of agreement and case marking that would correlate in the proper way with animacy, definiteness, and case. In line with the general observation that highly "grammatical" morphosyntactic properties (e.g., grammatical relations (DeLancey 1981, Croft 1983) and syntactic categories (Croft 1984)) actually represent pragmatic—specifically, discourse-functional—properties, namely the organization of information by the speaker for presentation to the hearer, the following definitions for agreement and case marking are proposed.

1.3.1 Agreement

Agreement—i.e., person-based agreement, also called "cross-reference" or "indexing"—indexes the *important* or *salient* arguments. This concept is a pragmatic one: salience is a relationship between the speaker and a referent in the described situation—that is, the speaker's attitude or point of view towards the referent—rather than a relation between

two entities in the described situation itself. Salience correlates with being high on the case, animacy and definiteness hierarchies, since the most salient entities are those most closely involved in the described event, closest in nature to the speaker, and most easily identifiable. The natural correlation predicts that where the presence vs. absence of agreement is grammaticalized, it will always align itself with high animacy, high definiteness, and core grammatical relations.

In addition to the evidence we have already presented, all of which is consistent with this definition, we can account for other typological phenomena. First, it is very frequently the case that the third person agreement "marker," which indexes the lower animacy (3rd person and nonpronominal) entities, is null for either subject or object. Among the many languages which have zero agreement in the third person are Quiché (singular direct object), Georgian (singular direct object; Comrie 1981:216), Yap (singular direct object; Jensen 1977), Gulf Arabic (masculine singular subject; Qafiseh 1977), Fula (singular subject; Arnott 1970), and Manam (some plural non-higher-animal direct and indirect object forms; Lichtenberk 1983).

Another very common phenomenon is that in languages with a two-argument agreement constraint, case roles such as recipient or experiencer which require mental capacity and therefore are occupied by humans (high in animacy and usually high in definiteness) are agreed with over case roles such as patient which are normally lower in animacy and/or definiteness. For example, in Kun-parlang, agreement is controlled by a hierarchy Recipient < Comitative < Patient. This hierarchy combines animacy and case: both recipients and comitatives are almost always humans, and the recipient (indirect object) case is higher on the case hierarchy than the comitative (oblique) case:

(8) nga- purrun- marnany- wom
1.SG.REAL- 3.DUAL- RECIP- return.PAST
'I returned to/for them (dual).'

(9) nga- purrum- walki- wom
1.SG.SBJ- 3.DUAL- COM- return.PAST
'I returned with them (dual).'

(10) *nga- marnany- purrun- walki- wom
1.SG.REAL- RECIP- 3.DUAL- COM return.PAST
'I returned to/for [no agreement] with them (dual)'

In Kinyarwanda (Kimenyi 1978) the verb agrees with patient possessors and benefactives even if instrumentals, locatives, manner, or directional arguments are promoted, and in most Bantu languages the dative/benefactive argument is obligatorily promoted (which triggers

agreement). There are also languages with a one-argument agreement constraint in which the verb agrees with whichever of subject and object is higher in animacy, namely, Chukchee and Tangut (Comrie 1980:231, 233).

Finally, Manam (Lichtenberk 1983) has a complicated set of rules determining agreement which combines the animacy hierarchy with the typically animate cases. The verb agrees with the dative recipient and not the patient if both are present; but the verb agrees with both the patient and the benefactive/source unless the patient is 1st or 2nd person, in which case the verb agrees with the patient only. Thus, the patient can trigger agreement over the typically animate benefactive only if the patient actually is high in animacy.

1.3.2 Case Marking

Case marking, in contrast to agreement, denotes *non-obvious* grammatical relations. This concept is also pragmatic. "Obviousness" does not denote a subclass of semantic relations between two entities. Rather, it denotes a relationship between the speaker (and hearer) and a semantic relation. A semantic relation between two entities may be obvious in certain discourse contexts but not in others, and this of course depends on the knowledge and presuppositions of the speaker and the hearer rather than on some property of the described situation. Case marking correlates with being low on the case hierarchy, since the relation of more oblique arguments to the predicate are less obvious than those of the central, normally present or even obligatory, arguments. Case marking also correlates with deviation from the natural correlations associated with a case position. When the animacy/definiteness properties associated with the case position are not the "natural" and thereby "obvious" ones, then the case position tends to be marked.

On the basis of the typological evidence, this definition predicts that the natural correlation of direct objects is with low animacy, low definiteness, and highly affected objects (i.e. genuine patients), and that the natural correlation of subjects is with high animacy and high definiteness, as well as high volitionality (as is generally considered to be the case):

	Subject correlations	*Object correlations*
Animacy	Human/Animate	Inanimate
Definiteness	Definite	Indefinite
Volitionality	Volitional	Affected

We have thus affirmed Comrie's suggestion that the unmarked direct objects are indefinite and inanimate—that is, direct objecthood correlates with low animacy and definiteness. The reason that Hopper and Thompson, who examined both agreement and case marking data,

considered highly individuated objects as indicators of Transitivity is that Transitivity—i.e., the morphosyntax that indicates Transitivity—correlates with a verb's having *two* distinct and highly salient arguments, not a surprising fact. Also, this analysis of the natural correlations of direct objects renders direct objects as exactly opposite to subjects, the other core argument position, and this is consistent with Rosch's contention that prototype categories tend to be as contrastive as possible with adjacent prototypes (Rosch 1978:37). In addition, by far the least marked direct objects are those which become incorporated into the verb, and those tend to be low in definiteness and in animacy (cf. Mithun 1984, especially p. 863; Sadock 1985). Finally, there is textual markedness evidence that suggests indefinite direct objects are unmarked: in an English text count, Givón (1979:51-52) found that 50% of the direct objects were indefinite and 82% of the indefinite NPs were direct objects.

In addition to the evidence presented so far, all of which is consistent with this definition, the definition will also account for a number of other typological facts. First, low animate/definite subjects as well as high animate/definite objects are sometimes case marked. This results in "classic" split animacy systems (Silverstein, 1976; the following examples are from Dixon 1979:87):

A	-∅	*-ngu*	-∅	[nasalization]	[nasalization]
S	-∅	-∅	-∅	-∅	-∅
O	*-na*	-∅	*-a*	*-a*	-∅
	1st & 2nd person pronouns	3rd person pronouns, all nouns	1st & 2nd person pronouns	3rd person pronouns	proper names, common nouns
	Dyirbal		**Cashinawa**		

The correlation of animacy can also extend to different classes of intransitive subjects in so-called "active" languages. For example, Northern Pomo displays a case marking system split roughly between human and non-human with an unmarked form for human agents and nonhuman patients, where both the agent and patient cases are used on a subclass of intransitives depending on the volitionality of the performer of the action (O'Connor and Caisse 1981:277-279). Thus, the unmarked case is used for human performers of volitional actions and nonhuman performers of nonvolitional actions, which are the natural correlations of animacy and volitionality.

A similar phenomenon is found with the causal opposite of volitionality, namely, affectedness. Direct objects which are less affected by the action have nonzero case marking. In the Russian examples in (11)–(12) (Moravcsik 1978:266), the bread in (12) is less affected by virtue of having only a part of it affected. In the well-known English examples in (13)–(14) (see S. Anderson 1970), the case-marked object is the less affected one.

(11) Peredajte mne xleb
pass.IMP 1.SG.DAT bread.(NOM)
'Pass me the bread.'

(12) Peredajte mne xleb -a
pass.IMP 1.SG.DAT bread -GEN
'Pass me some bread.'

(13) John shot Harry.

(14) John shot at Harry.

Finally, entities which fall into highly semantically-specific classes such as measure terms and deictic terms for places, directions and times, when used in case roles such as locative, allative, or extent which require equally specific semantic arguments, are not case-marked. This phenomenon is quite widespread, although the examples in (15)–(16) are from English:

(15) George Washington slept here/in this bed/*this bed.

(16) John ran five miles/across the field/*the field.

One also occasionally finds the converse phenomenon. In Malay (Dodds 1977:13), the prepositions *kĕ* 'to' and *dari* 'from' are used for motion to and from NPs which normally denote places. However, if the motion is to or from a person, an additional case marking is thus used for this, non-obvious, argument: the preposition *pada* 'at' must be added, so that the double prepositions *kĕpada* and *daripada* are used instead of the single ones.

2 Agreement, Case Marking, and Possession

If the definitions proposed for (person-based) agreement and case marking are indeed correct, then the correlations they predict should be valid in other grammatical domains in which the two relation-indicating strategies are used. There is one other domain in which both person-based agreement and case marking are used, namely, possession, and there is some evidence that the natural correlations do apply in this domain as well. The domain shift from verbal case to possession can be accomplished by mapping animacy and definiteness into themselves, and mapping the case hierarchy into alienability of possession, in which

inalienable possession is higher than alienable possession, obeying a hierarchy of: Body parts, Kinship, Part/Whole < Clothing, Tools < Other (cf. Seiler 1983).

There are some examples of properties correlating with cross-reference that control variation in the presence of cross-reference, although they are not numerous. The properties found are alienability, and animacy. For example, in Kanuri, a Nilo-Saharan language, the possessed item usually agrees with the possessor. If the possessor is postposed and the relation is alienable, however, then there is no agreement (Hutchison 1981:198-199; Moravcsik 1974:28 observed that preverbal position of objects correlates with presence of agreement, and postverbal with absence thereof):

(17) yâ -nzə́ álì -bè
mother -3.SG Ali -GEN
'Ali's mother'

(18) álì -bè fə̂r -nzə́
Ali -GEN horse -3.SG
'Ali's horse'

(19) fə̂r álì -bè
horse Ali -GEN
'Ali's horse'

In a number of Polynesian languages, possession is mediated by a so-called possessive particle or classifier. If the possession relation is inalienable, then a cross-reference marker is used, and if it is alienable, then the cross-reference marker is absent; the examples are from Hawaiian (Lichtenberk 1983b:162):

(20) k -o -na lima
ARTICLE -CLASS -3.SG hand
'his hand'

(21) nā kānaka o ke ali'i
ARTICLE people CLASS ARTICLE chief
'the people of the chief/the chief's people'

Finally, one finds zero third person singular possessive affixes, for example in Manam (Lichtenberk 1983a:264) and Dakota (Boas and Deloria 1941:127).

With case marking the evidence is more abundant. It is quite common to find case marking absent in inalienable possession relations and present in alienable possession relations; the following examples are

from Awa, a Papuan language (McKaughan 1973:22, 32; cf. also Dixon 1980:293):

(22) adena- (a)hde
2.SG- ear
'your ear'

(23) se -ne nah
3.PL -GEN house
'their house'

(24) iya -ne nah
dog -GEN house
'the dog's house'

Finally, English complex nominal constructions provide additional evidence that the same concept of "obviousness" is primary in noun-noun relations. The semantic relations which can be found in complex nominal constructions appear to vary in indefinitely many ways (Downing 1977). However, the relation between the two nouns must be pragmatically obvious, either through conventionalization in the case of grammaticized complex nominals such as *fire engine*, or through contextual factors in the case of innovations such as *grove map*. If neither of these conditions apply, then one must use an [N PP] paraphrase such as *map of memorial groves*, with an explicit case marking relating the two nouns. The pragmatic status rather than the semantics of the relation determines whether or not the complex nominal construction may be used.

3 Conclusion

The combination of typological analysis and a generalized marking theory which includes the notion of *relative markedness* and unifies markedness, hierarchies and prototypes, has allowed us to propose that the presence of (person-based) agreement and case marking each have a pragmatic significance which is universal, although their use varies across languages and is often grammaticalized. The next step is to seek an explanation for the pragmatic significance: why does agreement index the important arguments and case marking indicate the non-obvious relations?

The explanation can be found in the different ways in which agreement and case marking serve the same function, namely, to express a relation between two entities. Case marking is a *relational* strategy: the case marker denotes the relation that holds between the two entities. Agreement is a *deictic* strategy: the agreement marker actually denotes

the other entity that is related to the entity denoted by the agreeing constituent. This is a *semantic* characterization of the two strategies, since the definitions are based on the denotations of the morphemes involved. A semantic definition is required because the typological generalizations hold regardless of whether the morpheme involved is a particle, an affix, or even an internal morphophonemic alternation, and whether the morpheme is associated with the head or the dependent.[6] Our hypothesis is, however, that discourse-functional factors (salience and obviousness) determine the *presence* of either strategy in the surface structure; we assume that the semantics and the pragmatics interact closely.

Case marking is a complement of the strategy of simple juxtaposition of the related constituents, in which the hearer must infer the relation that holds between the two. Simple juxtaposition is only possible when the relation between the two terms is obvious enough for the hearer to easily infer it. Otherwise, the relation must be more explicitly represented in the utterance, and case marking is the strategy for doing so. This tends to be with the more peripheral and less prototypical participants. This appears to true no matter where the relation-indicating morpheme is located syntactically. In all of our examples so far, the case marker we have examined is either an affix on the constituent denoting

[6] The head- versus dependent-marking distinction figures importantly in a recently published paper by Johanna Nichols (Nichols 1986). Although there is no space to comment extensively on Nichols' arguments, the following remarks support the preference of the deictic/relational distinction over the head-/dependent-marking distinction. First, the head-/dependent-marking distinction cannot incorporate independent elements, which Nichols must describe as "neutral"; yet adpositions clearly fall under the same generalizations as case markers. (These generalizations apply to oblique arguments as well, which Nichols excludes from her analysis.) Second, the head-/dependent-marking distinction divides person-based verbal/possessive agreement and adjectival agreement, conflating the latter with case marking; yet adjectival agreement has much more in common with person-based agreement, both being deictic strategies, than with case marking. Finally, the basic head/dependent distribution of deictic and relational strategies can be explained on independent grounds. Verbs, adjectives, and adpositions are *inherently relational*, that is, they conceptually require additional entities (their "arguments"). Nouns are not inherently relational, except for "relational nouns" such as body parts and kinship terms, and except when they are functioning as predicates (Croft 1984). The relational lexical items must somehow point to the fillers of their "arguments" by indicating properties of the fillers (person, number, gender); hence the use of the deictic strategy on predicates (heads) in general, adjectives (dependents) and typically relational possessive heads. On the other hand, the nonrelational lexical items must indicate what their relation to the head is, since that information is not inherently present in the nonrelational item itself; hence the use of the relational strategy with nominal dependents.

the dependent entity or an adposition governing that entity. A relation-indicating morpheme may also appear on the constituent denoting the head entity, in which case it is called an applicative affix. These appear to be historically related to adpositions, and sometimes the morpheme may be associated with either the head or the dependent constituent, as in Mokilese (Harrison 1976:163-164), Abkhaz (Hewitt 1979:113-114) and Kun-parlang. Even in these cases, however, the morpheme obeys our predictions: for example, it is always an oblique relation such as benefactive, locative or instrumental which requires a nonzero applicative morpheme, not the patient (the normal direct object).

The deictic strategy appears to be a strategy of *person* deixis. As we have noted, the generalizations we have proposed do not appear to apply to agreement within NP's, normally based on *gender*, and only applies to agreement with dependent NP's. In fact, the natural correlations we have observed appear to be valid not only for "true" agreement—in which the deixis is endophoric—but also for agreement systems which appear to be fused pronominals, where the domain of agreement is exactly complementary to the domain of independent NP arguments (S. Anderson 1982:579; Mithun 1986). They even hold for pronominals not morphologically bound to the head, such as for example the object clitics in Hausa. The Hausa object clitics are optional when the direct object is specific, but prohibited when it is not (Cowan and Schuh 1976:135). They are also required with *sanì* 'know' and *ganī* 'see' when the direct object is a concrete, physical object and prohibited when the direct object is an abstract object or an activity, and can be used with the semantic recipient of a small set of verbs, including the verb *bā* 'give' (Cowan and Schuh 1976:135-137). All of these constraints conform with our hypothesis.

Thus the explanation for the meaning of the deictic strategy as indicating *salient* referents must be sought in the nature of personal pronominal reference. Both pronouns and agreement markers are used to identify and maintain the identity of their referents across the discourse (see DuBois 1980; Lehmann, this volume). There are certainly processing constraints on how much cross-referencing of entities can be handled at once by a person—this is manifested, for example, in the common two-argument agreement constraint discussed above. Therefore, the speaker must make a choice as to which entities will continue to be cross-referenced and which ones will not. Naturally, the most important or salient entities will continue to be cross-referenced, and those tend to be the most animate ones, the most definite ones, and the ones most central to the events being reported.

Finally, the deictic and relational strategies must be situated in the context of possible strategies for expressing relationship between entities in discourse. As we have already pointed out, the person-based strategy whose behavior we have examined is only one of at least two

types of deictic strategies, and the relational strategy is the complement of the strategy of simple juxtaposition of constituents, the latter being the "null" relational strategy. The deictic and relational strategies exhaust the morphological possibilities for expressing a relation between two entities, since the morpheme involved must denote either the relation itself or one of the elements. The only other possible strategy for relating entities is the syntactic one of word order. Word order appears to be independent of the other strategies, which are defined in terms of morphemes (case vs. agreement) and morpheme denotations, and it appears to have its own organizing principles, such as "attention flow" (DeLancey 1981) or "newsworthiness" (Mithun, to appear).

In closing, we may point out that the deictic-relational distinction in morphosyntactic strategies for indicating grammatical relations is grammatically significant at all levels of syntactic structure, not just the clausal level that we have described here in detail. At the phrasal level we find the adnominal modification discussed in Section 3, and also the deictic but gender-based agreement, and the (probably relational) linking particles of Persian and Austronesian. Lichtenberk (1983b) argues for a distinction between the deictic numeral classifiers and the relational possessive classifiers. Finally, at the sentential level we find connectives and subordinators which are both deictic and relational in historical origin and, we expect, in grammatical behavior.

References

Anderson, A. 1973. *Rules of the Aztec Language.* Salt Lake City: University of Utah Press.

Anderson, S. R. 1970. A little light on the role of deep structure in semantic interpretation. NSF Report No. 26, II.1–II.13. Cambridge, Mass.: Harvard University.

Anderson, S. R. 1982. Where's morphology? *Linguistic Inquiry* 13:571–612.

Arnott, D. W. 1970. *The Nominal and Verbal Systems of Fula.* Cambridge, England: Cambridge University Press.

Ashton, E. O. 1944. *Swahili Grammar.* London: Longman.

Boas, F. and E. Deloria. *Dakota Grammar.* (Memoirs of the National Academy of Sciences, Vol. 23, 2d memoir.) Washington: U.S. Government Printing Office.

Clark, H. D. B. and M. Hamamura. 1981. *Colloquial Japanese.* London: Routledge and Kegan Paul.

Comrie, B. 1976. The syntax of causative constructions: cross-linguistic similarities and differences. In M. Shibatani (Ed), *The Syntax of Causative Constructions.* (Syntax and Semantics, Vol. 6.) New York: Academic Press.

Comrie, B. 1979. Definite and animate direct objects: a natural class. *Linguistica Silesiana* 3:13–21.

Comrie, B. 1980. Agreement, animacy and voice. In G. Brettschneider and C. Lehmann (Eds), *Wege zur Universalienforschung,* 229–234. Tübingen: Gunter Narr.

Comrie, B. 1981. *The Languages of the Soviet Union.* Cambridge, England: Cambridge University Press.

Cowan, J. R. and R. G. Schuh. 1976. *Spoken Hausa.* Ithaca: Spoken Language Services.

Croft, W. 1983. Grammatical relations vs. thematic roles as universals. *Papers from the Nineteenth Regional Meeting of the Chicago Linguistic Society,* 76–94.

Croft, W. 1984. Semantic and pragmatic correlates to syntactic categories. *Papers from the Parasession on Lexical Semantics, Twentieth Regional Meeting of the Chicago Linguistic Society,* 53–71.

DeLancey, S. 1981. An interpretation of split ergativity and related patterns. *Language* 57:626–657.

Dixon, R. M. W. 1972. *The Dyirbal Language of North Queensland.* Cambridge, England: Cambridge University Press.

Dixon, R. M. W. 1979. Ergativity. *Language* 55:59–138.

Dixon, R. M. W. 1980. *The Languages of Australia.* Cambridge, England: Cambridge University Press.

Dodds, R. W. 1977. *Teach Yourself Malay.* Sevenoaks, Kent: Hodder and Stoughton.

Downing, P. 1977. On the creation and use of English compound nominals. *Language* 53:810–842.

Dryer, M. S. 1983. Indirect objects in Kinyarwanda revisited. In D. Perlmutter (Ed), *Studies in Relational Grammar* 1, 129–140. Chicago: The University of Chicago Press.

DuBois, J. 1980. Beyond definiteness: the trace of identity in discourse. In W. Chafe (Ed), *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production.* Norwood, N. J.: Ablex.

Givón, T. 1976. Topic, pronoun and grammatical agreement. In C. Li (Ed), *Subject and Topic,* 149–188. New York: Academic Press.

Givón, T. 1979. *On Understanding Grammar.* New York: Academic Press.

Greenberg, J. H. 1966. *Language Universals, With Special Reference to Feature Hierarchies.* (Janua Linguarum Series Minor, 59.) The Hague: Mouton.

Greenberg, J. H. 1978. How does a language acquire gender markers? In J. H. Greenberg (Ed), *Universals of Human Language* 3, 47–82. Stanford: Stanford University Press.

Harrison, S. P. 1976. *Mokilese Reference Grammar.* Honolulu: University Press of Hawaii.

Hewitt, B. G. 1979. *Abkhaz.* (Lingua Descriptive Series, 2.) Amsterdam: North Holland.

Hopper, P. and S. Thompson 1980. Transitivity in grammar and discourse. *Language* 56:251–299.

Hutchison, J. P. 1981. *The Kanuri Language.* Madison: African Studies Program.

Jensen, J. T. 1977. *Yapese Reference Grammar.* Honolulu: University Press of Hawaii.

Keenan, E. and B. Comrie 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8:63–99.

Kimenyi, A. 1978. *A Relational Grammar of Kinyarwanda.* (University of California Publications in Linguistics, Vol. 91.) Berkeley: University of California Press.

Ladziņa, T. B. 1966. *Teach Yourself Latvian.* London: English Universities Press.

Lewis, G. L. 1967. *Turkish Grammar.* Oxford: Oxford University Press.

Lichtenberk, F. 1983a. *A Grammar of Manam.* (Oceanic Linguistics Special Publication No. 18.) Honolulu: University Press of Hawaii.

Lichtenberk, F. 1983b. Relational classifiers. *Lingua* 60:147–176.

Mace, J. 1962. *Teach Yourself Persian.* London: English Universities Press.

Matisoff, J. 1973. *The Grammar of Lahu.* (University of California Publications in Linguistics, Vol. 75.) Berkeley: University of California Press.

McKaughan, H. (Ed) 1973. *The Languages of the Eastern Family of the East New Guinea Highlands Stock.* Seattle: University of Washington Press.

Miller, W. R. 1965. *Acoma Grammar and Texts.* (University of California Publications in Linguistics, Vol. 40.) Berkeley: University of California Press.

Mithun, M. 1984. The evolution of noun incorporation. *Language* 60:847–894.

Mithun, M. 1986. When zero isn't there. Paper presented at the 12th Annual Meeting of the Berkeley Linguistics Society.

Mithun, M. To appear. Is basic word order universal? In R. Tomlin (Ed), *Discourse Relations and Cognitive Units.*

Moravcsik, E. 1974. Object-verb agreement. *Working Papers in Language Universals* 15:25–40. Stanford: Department of Linguistics, Stanford University.

Moravcsik, E. 1978. On the case marking of objects. In J. H. Greenberg (Ed), *Universals of Human Language* 4, 249–290. Stanford: Stanford University Press.

Nichols, J. 1986. Head-marking and dependent-marking grammar. *Language* 62:56–119.

O'Connor, C. and M. Caisse. 1981. Aspects of role-dominance in Northern Pomo. *Papers from the Seventeenth Regional Meeting of the Chicago Linguistic Society,* 277–285.

Parker, G. 1969. *Ayacucho Quechua Grammar and Dictionary.* (Janua Linguarum Series Practica, 82.) The Hague: Mouton.

Perrott, D. V. 1972. *Teach Yourself Swahili.* New York: David McKay Company, Inc.

Qafiseh, H. 1977. *A Short Reference Grammar of Gulf Arabic.* Tucson: University of Arizona Press.

Rosch, E. 1978. Principles of categorization. In E. Rosch and B. Lloyd (Eds), *Cognition and Categorization,* 27–48. Hillsdale, N. J.: Lawrence Erlbaum Associates.

Sadock, J. M. 1985. The Southern Tiwa incorporability hierarchy. *International Journal of American Linguistics* 51:568–572.

Seiler, H. 1983. *Possession as an Operational Dimension of Language.* Tübingen: Gunter Narr.

Shackle, C. 1972. *Teach Yourself Punjabi.* London: English Universities Press.

Silverstein, M. 1976. Hierarchy of features and ergativity. In R. M. W. Dixon (Ed), *Grammatical Categories in Australian Languages,* 112–171. Canberra: Australian Institute of Aboriginal Studies.

Thomas, D. D. 1971. *Chrau Grammar.* (Oceanic Linguistics Special Publication No. 7.) Honolulu: University Press of Hawaii.

Whitney, A. 1944. *Colloquial Hungarian.* London: Routledge and Kegan Paul.