

CONFERENCIAS PLENARIAS

EXPLANATION IN LANGUAGE TYPOLOGY*

Bernard Comrie

University of Southern California

Introduction

After a period - the heyday of American Structuralism - when it was virtually forbidden to speak approvingly of language universals, there have more recently been two main approaches to the problem of language universals, and the related problem of typology, within American linguistics. First, the variety of transformational-generative grammar that has grown around the works of Noam Chomsky expresses an explicit interest in the search for language universals; second, linguists working with or influenced by Joseph Greenberg have adopted a rather different conception of language universals, and a very different methodological approach. The main differences between these two approaches can be characterized as follows. Chomsky argues that the search for language universals is best carried out on the basis of the detailed study of a single language, and that language universals are most fruitfully predicated of abstract structures and rule systems. Greenberg, however, argues that in order to come up with potentially valid language universals it is necessary to work with a wide range of languages, and moreover that fruitful universals can be predicated of levels of analysis much closer to the surface.

There is one further point where the classical Chomskyan and Greenbergian positions can be contrasted. From very early on in his work Chomsky has been interested in explaining language universals, usually by appealing to the child's innate language-learning potential - although in the absence of independent evidence concerning this language-learning potential, its validity as an explanation is seriously compromised. Within the Greenbergian school, however, much greater emphasis has been placed on establishing a number of universals, i.e. in establishing a sufficient database so that, later on, it might be possible to seek explanations, but for generalizations that are already firmly established. On most of the points where the Chomskyan and Greenbergian approaches differ, I am in agreement with Greenberg. However, it seems to me that our knowledge of language universals and language typology has already progressed to a stage where we can start constructing explanations for some, at least, of the observed regularities across languages, with good possibilities for evaluation of the explanations we posit. This is the main point I want to illustrate in this contribution.

Before looking at some actual data, it is worth pausing for a moment to make more explicit what is meant by 'explanation' in this context. I shall assume that explanation consists in showing that apparently unrelated phenomena are in fact related. While this might not seem to correspond to one's immediate intuition of what explanation is, I would argue that it does correspond to the use of the term explanation in practice - both in the practice of scientists and in the practice of ordinary language use. Note that as such, explanation is in a sense a never-ending process: once we have shown that two phenomena, hitherto thought to be unrelated, are in fact related, we are still faced with the problem of relating this now uniform phenomenon to other phenomena. But this is a general characteristic of human investigations: finding one generalization only spurs the investigator on to seek for even more general formulations.

On a more specific note, I will be arguing that in many instances, at least, one can demonstrate that certain typological regularities of language structure can be related to extralinguistic, or at least extrastructural (pragmatic) regularities. This differs from the Chomskyan view that explanations for structural regularities are to be formulated in intralinguistic terms, as formal statements about language structure. Or at least, there is a difference in emphasis here: I do not deny that there may be structural regularities that have no explanation beyond language structure, just as Chomsky has never denied that certain aspects of language structure may have extralinguistic explanations. However, it seems to me that concentration on intralinguistic explanations, especially where there are no independent motivations, tends to preclude the possibility of seeing extralinguistic explanations; and since an extralinguistic explanation serves to relate apparently more disparate phenomena than does an intralinguistic explanation, the former leads to a much more encompassing overall explanatory theory.

The discussion so far has been rather abstract, and it is time now to illustrate these general points with discussion of some actual data. Rather than review a wide range of data areas, with correspondingly superficial discussion of each, I have chosen instead to look at one data area in much greater detail, namely the general problem of ergativity. For a more extensive general survey of ergativity, reference may be made to Comrie (1978) or Dixon (1989); the account in section 1 is necessarily much abbreviated.

1. Degrees of Ergativity

Discussion of ergativity requires a distinction to be made between one-place (intransitive) predicates, such as come, and two-place (transitive) predicates, such as hit. Moreover, we must have a way of identifying the arguments (accompanying noun phrases) of a predicate that remains constant across languages with different morphological and syntactic types. I shall use the symbol S to identify the single argument of a one-place predicate, and the symbols A and P to identify the two arguments of a two-place predicate, A being the argument which is most typically (though

not invariably, for all two-place predicates) a semantic agent and P the argument which is most typically a semantic patient. The symbols S, A, and P are clearly mnemonic for 'subject', 'agent', and 'patient', but are not identical with these syntactic and semantic relations; to avoid confusion, they are best treated as arbitrary symbols.

In most, if not all, languages, some distinction, morphological or syntactic, is made between the A and the P of the transitive construction. Taking morphology as the clearest manifestation of this opposition, many languages have different cases for A and P in the transitive construction, as in the following examples, from English and the Australian Aboriginal language Dyirbal (Dixon 1972; all Dyirbal material below is taken or adapted from this source):

1. he (NOM) hit me (ACC).
2. balan djugumbil baḡḡul yaṛaḡgu balgan.
CLASS-ABS woman-ABS CLASS-ERG man-ERG hit
'the man hit the woman.'

It is also usually the case that, where such a distinction is made between A and P in the transitive construction, the S of the intransitive construction is identified with either the A or the P of the transitive construction, rather than being treated as a third entity. (Note that A and P contrast in the transitive construction, whereas there can be no contrast between S and A or between S and P - they do not occur in the same construction.) In English, for morphological purposes S is identified with A, as in (3):

3. he (NOM) came.

This is the so-called nominative-accusative system, with S and A (nominative) treated alike, and P (accusative) treated differently. In Dyirbal, however, for (most) morphological purposes S is identified with P, as in (4):

4. balan djugumbil baninju.
CLASS-ABS woman-ABS came
'the woman came.'

This is the so-called ergative-absolutive system, with S and P (absolutive) treated alike, and A (ergative) treated differently.

Although we have illustrated the difference between nominative-accusative and ergative-absolutive types by means of noun phrase morphology, the same pattern can also be realized in syntax. For instance, English has a constraint on coordination with omission of the coreferential noun phrase in the second conjunct whereby the coreferential noun phrases must both be either S or A of their clause, not P; this constraint is to be interpreted as allowing that the noun phrase might be S of one clause and A of the other. Thus (5) can only be interpreted as the coordination of (6)

and (7), not as the coordination of (6) and (8):

5. the man hit the woman and came here.
6. the man hit the woman.
7. the man came here.
8. the woman came here.

Dyirbal has a corresponding constraint, but in Dyirbal the coreferential noun phrases must be either S or P, not A, of their clauses. Thus (9) can only be interpreted as the coordination of (10) and (11), not as the coordination of (10) and (12):

9. balan djugumbil (ABS) baḡgul yaṛaṅgu (ERG) balgan, baninju.
'the man hit the woman, and she came here.'
10. balan djugumbil (ABS) baḡgul yaṛaṅgu (ERG) balgan. (= (2))
'the man hit the woman.'
11. balan djugumbil (ABS) baninju. (= (4))
'the woman came here.'
12. bayi yaṛa (ABS) baninju.
'the man came here.'

English and Dyirbal happen to be languages whose morphology and syntax is almost consistently nominative-accusative and ergative-absolutive, respectively. However, there are also languages which combine nominative-accusative and ergative-absolutive systems to different degrees, in different combinations, in their overall structure. One such language is Chukchee, a Paleosiberian spoken in the far north-east of Siberia; the discussion below follows Comrie (1979). The morphology of noun phrases in Chukchee works very consistently on an ergative-absolutive basis, as can be seen from (13)-(14):

13. reqokalḡan yetḡi.
fox-ABS came
'the fox came.'
14. riḡukete ḡenulin teḡiḡḡan.
fox-ERG ate meat-ABS
'the fox ate the meat.'

However, there are other areas where Chukchee works on the nominative-accusative system, for instance in the syntax of infinitival constructions. In examples (15) and (16), the infinitival 'clause' has been marked off by square brackets in both the Chukchee sentences and the English glosses:

15. gəmnan gət tite məwinretgət [ermetwik].
 I-ERG you-ABS some-time may-I-help-you to-get-strong
 'let me help you some time [to grow strong].'
16. morgənan gət mətrewinretgət [riwlək əməlʔo
 we-ERG you-ABS we-will-help-you to-move all
 gečeyot].
 collected-items-ABS
 'we will help you [to move all the collected items].'

The English examples illustrate a general feature of English syntax, namely that both S and A (but not P) can be omitted in infinitive clauses under coreference with a specified noun phrase in the main clause: in (15) I am to help you so that you (S) may grow strong; in (16), we will help you so that you (A) may move the collected items. The Chukchee sentences evince precisely the same distribution, i.e. it is the S or A of a Chukchee infinitive that may be omitted: this is the nominative-accusative system, just as for the parallel construction in English, even though Chukchee is ergative-absolutive in its noun phrase morphology.

Work similar to that just discussed has established that different languages are ergative (exhibit 'split ergativity') to different degrees, i.e. are characterized to different extents by the combination of ergative-absolutive and nominative-accusative systems. The precise combinations of these two systems thus provide a typological parameter, enabling us to typologize languages according to how much ergativity they have, and where precisely this ergativity manifests itself. However, it turns out that we can go even further than this in a combined study of typology and universals. Our discussion so far would permit a language to manifest the ergative-absolutive system anywhere in its morphology and syntax, and likewise to manifest the nominative-accusative system anywhere in its morphology and syntax. Detailed investigation of ergativity in a wide range of languages suggests rather that there are certain principles governing the distribution of these two systems.

For instance, there are certain constructions that have a much greater tendency to operate on the nominative-accusative system; in section 2, I discuss omission of the addressee in imperatives as one such construction. Conversely, there are certain constructions that tend to gravitate towards the ergative-absolutive system, and in section 3 I discuss resultative constructions as an example of this set. Note that I am not claiming that there are certain constructions that are necessarily nominative-accusative, or necessarily ergative-absolutive (though this possibility is not excluded), rather I am claiming that certain constructions have a strong tendency, in languages of different genetic families and from different parts of the world, to gravitate towards one or other of ergative-absolutive and nominative-accusative. Thus the distribution of constructions on this parameter is not random. Having established this, we would like also to find an explanation for this distribution, and

in sections 2 and 3 I will suggest that this explanation is to be sought by relating formal aspects of syntactic structure (ergative-absolutive versus nominative-accusative systems) to pragmatic considerations.

2. Imperatives and the Nominative-Accusative System

In his general study of ergativity, Dixon (1979:112-114) observes that many languages allow the possibility of omitting overt reference to the addressee in second-person imperatives. In English, for instance, the imperatives come here and hit the man are possible, indeed strongly preferred, alongside the more explicit you come here and you hit the man. In English, as these examples illustrate, the omitted addressee noun phrase may be either S of the intransitive construction or A of the transitive construction, whereas parallel omission of P of the transitive construction is impossible, i.e. there is no imperative of the form *the man hit meaning 'may the man hit you'. Thus addressee omission in the English imperative is an instance of the nominative-accusative system in syntax.

Since English is almost exclusively nominative-accusative in its syntactic structure, this result is not surprising even against the background of a purely formal approach to English syntax. What is surprising from such a viewpoint, however, is the fact that Dyrbal, despite its near-total adherence to the ergative-absolutive system in syntax, evinces precisely the same distribution as in English in imperative constructions. In such constructions, it is possible (though not obligatory) to delete a second person S or A, as in (17) and (18):

17. (ŋinda) bani.¹
you-NOM come
'(you) come here.'
18. (ŋinda) bayi yara balga.
you-NOM CLASS-ABS man-ABS hit
'(you) hit the man.'

It is not possible to omit a second-person P in a transitive imperative construction in this way.

Following the approach advocated by Dixon in his discussion of such examples, we may argue that is a principled reason for the apparently exceptional behaviour of addressee omission in Dyrbal imperatives. For an imperative to encode a felicitous speech act, it is necessary that the person to whom the imperative is addressed should have, in principle, the power to carry out the instruction in question, i.e. should be high in agentivity or control. In the most frequently occurring kind of imperative - in many languages, the only kind -, the person to whom the instruction is given is necessarily the addressee (second person). There is thus a high correlation, in a felicitous imperative, between addressee and participant with high

degree of control. If we now ask about the syntactic correlates of 'participant with high degree of control', then we note that in the transitive construction the A almost invariably has a higher degree of control than the P, e.g. with the verb hit the person who hits exercises more control than the person who gets hit. In intransitive constructions, to the extent that control is involved (i.e. with verbs like come expressing voluntary actions), control lies with the S. In imperative constructions, therefore, there are pragmatic reasons for expecting S and A to behave alike. Indeed, one can take the argument a stage further. Since most typical use of imperatives is to direct a command to an addressee who must have control over the situation the speaker wants to come about, allowing addressee omission to apply to S and A is an economical way of allowing the formally most concise construction to refer directly to the most typical use of imperatives. The explanation could therefore be referred to as functional-pragmatic; it has no basis in the formal structure of Dyrbal, indeed the observed distribution goes against the usual ergative-absolutive syntax of Dyrbal.

If this explanation is correct, then one would expect the same distribution to be found with indirect commands. And indeed, McKay (1976:503-504) notes that in Rembarnga, an Australian Aboriginal language with widespread syntactic ergativity, indirect command constructions permit deletion of the argument of the embedded clause, coreferential with the object of the main clause (recipient of the command) if the argument in question is S or A, but not P:

19. ɲanpayinəʒwa roɲəkan.
 they-told-me to-go
 'they told me to go.'
20. ɲinjyinaʒwa ɾe teʒwanəkan.
 I-told-you meat to-give
 'I told you to give the meat.'

In (19), the S of 'to go', i.e. you, is omitted; in (20), the A of 'to give', i.e. me, is omitted.² Compare also Chukchee sentences (15)-(16) above.

3. Resultatives and the Ergative-Absolutive System

Although it has been quite widely recognized in linguistic literature that there are semantic-pragmatic reasons for grouping S and A together in terms of agentivity or control, there is also some less widely recognized evidence in favor of a semantic-pragmatic grouping together of S and P, which finds syntactic realization in a bias towards ergative syntax in certain constructions. In this section, I will first present some data, using the Paleosiberian language-isolate Nivkh (Gilyak) as the source of the illustrative material (Nedjalkov et al. 1974), turning subsequently to possible explanations for the distribution found - a more detailed account of the explanatory principles involved may be found in Comrie (MS).

In Nivkh syntax overall, there is considerable use of the grouping of S and A together as opposed to P, i.e. of nominative-accusative syntax. One instance of this which will play a role in the ensuing discussion is that P, but not S or A, conditions certain morphophonemic changes in the initial consonant of a verb. For example, the transitive verb 'to roast' has the citation form $\check{y}ad\bar{}$ (where \bar{d} is a finite verb suffix). In sentence (21), however, this verb appears as $t^had\bar{}$, the change to initial t^h being conditioned by the final s of the P $\check{E}us$ 'meat':

21. $umgu \check{E}us t^had\bar{}$
 woman meat roast
 'the woman roasted the meat.'

The Nivkh construction in which we are particularly interested is the resultative construction, in which the verb takes the suffix $-xətə$; the meaning of this construction is that a certain state holds as the result of a previous event. This can be seen in its simplest form with intransitive verbs, as in (22)-(23):

22. $anaq yod\bar{}$
 iron rust
 'the iron rusted.'
23. $anaq yoxətad\bar{}$
 iron rust-RESULT
 'the iron has rusted.'

Sentence (22) simply refers to the event of the iron's turning rusty, an event that may since have been negated by removing the rust; sentence (23), on the other hand, refers directly to the present state of the iron's being rusty, this state being the result of the event of its turning rusty.

In general, resultatives of transitive verbs, unlike nonresultative forms of such verbs, allow only a single argument, i.e. one of the arguments of the two-place transitive verb must be omitted. For intransitive verbs, we have seen that the S of the nonresultative appears as the sole argument (S) of the resultative. What happens, then, in the resultative of a transitive construction like (21)? The answer is that the P remains as sole argument (S) of the resultative form, as in (24):

24. $tūs \check{y}axətad\bar{}$
 meat roast-RESULT
 'the meat has roasted.'

In (24), since $\check{E}us$ is no longer a P, the initial consonant of the verb remains as in its citation form. Note, incidentally, that (24) means, as indicated, 'the meat has been roasted' (sc. by someone unspecified), and not 'the meat has roasted' (sc. conceivable by spontaneous means), i.e. we are dealing with the transitive verb 'to roast'

rather than with its corresponding intransitive lexical counterpart. The phenomenon illustrated is thus that the Nivkh resultative involves attribution of a state (as the result of a change of state) to the S of an intransitive verb or to the P of a transitive verb: this like treatment of S and P, in contrast to A, is an instance of ergative-absolutive syntax, despite the generally nominative-accusative nature of Nivkh syntax.

The motivation for ergative syntax here can be uncovered if we examine the semantics and pragmatics of resultative constructions, which in a wide range of languages involve bias towards ergativity. The resultative construction attributes a certain state, as the result of a change of state, to a certain entity. With one-place predicates, the change of state is naturally attributed to the only participant, the S. With two-place predicates, however, it is in the vast majority of cases more natural to attribute the new state to the P, rather than to the A, of the predicate in question. For instance, John has killed Bill, has broken the vase, has melted the ice describe a change of state not, or at least not primarily, in John, but rather in Bill, the vase, or the ice. The only difference between English and Nivkh here is that in Nivkh this pragmatic property of resultative constructions has been grammaticalized, integrated into the syntax of the language in the form of syntactic ergativity.

There are of course some two-place predicates where there is either no change of state, or even attribution of the change of state to the A rather than to the P, as in John has reached the summit, has found out the information. In some languages with syntactic ergativity in resultative constructions, such sentences are simply assimilated to the general class of two-place predicates in their syntactic behaviour. Nivkh, however, provides evidence for the importance of the pragmatic explanation by not generalizing ergative syntactic behaviour where there is no change of state in the P, but rather in the A, so that in (25) and (26) the verb yimd 'find out' remains as a two-place predicates, with the same arguments in both nonresultative and resultative:

25. andx p^hrənəd̄ if yimd̄.
 guest will-come he find-out
 'he found out that the guest will come.'
26. andx p^hrənəd̄ if yimɣətad̄.
 guest will-come he find-out-RESULT
 'he has found out that the guest will come.'

4. Conclusions

Although, within the format of this presentation, I have been able to give illustrative material from only a very small number of languages, the kinds of phenomena discussed are in fact found in a wide range of languages, and it is only on the basis of this recurrence across a wide range of languages that I would feel confident in claiming that there is a bias towards nominative-accusative syntax in in-

peratives or towards ergative-absolutive syntax in resultative constructions. To this extent, the position adopted here is closer to the Greenbergian tradition than to the Chomskyan tradition of work on language universals. Another similarity to the Greenbergian tradition is that language universals are predicated not of abstract structures or rule systems, but rather of structures much closer to the surface: to surface syntactic structures, semantic interpretations, and to relations among these.

The search for explanations for language universals might seem to link the work outlined in this presentation to that of Chomsky rather than to that of Greenberg and his closest coworkers, and certainly one of the major contributions Chomsky has made to linguistics has been to make linguists interested in the search for explanatory principles. However, the precise kind of explanation advocated here falls well outside that envisaged in mainstream transformational-generative syntax. In particular, I have argued that explanations for syntactic phenomena are to be sought not solely within syntax, but that many patterns which seem quite arbitrary from a syntactic viewpoint find a natural explanation in terms of their pragmatic correlates.

NOTAS:

Abbreviations Used

ABS	Absolutive
ACC	Accusative
CLASS	Classifier
ERG	Ergative
NOM	Nominative
RESULT	Resultative

* This article is a revised version of my talk 'Syntactic Typology' given to the Programa Interamericano de Lingüística e Ensino de Línguas at the Universidade Estadual de Campinas, in January 1980. I am grateful to those who participated in the ensuing discussion, and also to my many other colleagues, too numerous to mention by name, who have contributed to the crystallization of the ideas contained in this paper.

1. The verb forms in (17) and (18) are imperatives, formed in Dyrbal by deleting the final consonant of the verb stem (baniy- 'come', balgal- 'hit'). The nonfuture tense, illustrated in earlier examples, involves replacement of stemfinal y by -nju and of l by -n. Pronouns in Dyrbal have a nominative-accusative (or, for some pronouns, mixed) case-marking system, whence nominative ninda in both (17) and (18); the accusative is ninuna; however, Dixon shows that the general syntactic behaviour of pronouns, as of nouns, is according to the ergative-absolutive system, as in (i):

- (i) *ŋayguna ŋinda balgan, baninju.*
 I-ACC you-NOM hit came
 'you hit me, and I came here.'

2. Dixon (1979:112) claims that in imperatives, S and A will necessarily be treated alike, at least for certain syntactic processes ('even the most ergative language will treat S and A NP's of imperatives the same'). I am not committed to this strong a position, although I have no counterexamples for direct imperatives, but would argue rather that the pragmatics pushes strongly in the direction of S-A identification rather than S-P identification. With regard to indirect commands in Dyirbal, Dixon (1979:128-129) himself notes that Dyirbal cannot directly delete the A of a transitive purposive complement (roughly paralleling the English infinitive), but must use a derived intransitive construction, in which the noun phrase in question is expressed as an S. Thus S, but not A, would be deletable in such constructions. Dixon considers that this is surprising, but argues further that the constraint is on underlying rather than surface grammatical relations, i.e. for this purpose the S of a derived intransitive would be treated as an A. In personal communication, Dixon has suggested that even this position may not be tenable, the crucial evidence being whether or not P can be deleted to give sentences meaning 'the woman told me that the man should hit me', i.e. 'the woman told me to get hit by the man.' Although such examples are not attested in Dyirbal texts, their grammaticality status remains, for the moment, open. If they are grammatical, then in the case of indirect commands Dyirbal would evince syntactic ergativity, despite the pragmatic pressure for identification of S and A.

REFERENCES:

- COMRIE, Bernard (1978). 'Ergativity'. In *Syntactic Typology: Studies in the phenomenology of Language*, ed. Winfred P. Lehmann. Austin: University of Texas Press. Pages 329-394.
- , (1979). 'Degrees of Ergativity: Some Chukchee Evidence'. In *Ergativity: Towards a Theory of Grammatical Relations*, ed. Frans Plank. London: Academic Press. Pages 219-240.
- , (MS). 'Aspect and Voice: Some Reflections on Perfect and Passive'. Department of Linguistics, University of Southern California, Los Angeles.
- DIXON, R.M.W. (1972). *The Dyirbal Language of North Queensland*. Cambridge University Press.
- , (1979). 'Ergativity'. *Language* 55, 59-138.

- McKAY, G.R. (1976). 'Rembarnga'. In *Grammatical Categories in Australian Languages*, ed. R.M.W. Dixon. Canberra: Australian Institute of Aboriginal Studies. Pages 494-505.
- NEDJALKOV, V.P., G.A. Otaina, & A.A. Xolodovič (1974). 'Diatezy i zalogi v nivxskom jazyke' ('Diatheses and Voices in the Nivkh Language'). In *Tipologija passivnyx konstrukcij: diatezy i zalogi (The Typology of Passive Constructions: Diatheses and Voices)*, ed. A.A. Xolodovič. Leningrad: Izd-vo 'Nauka'. Pages 232-251.