THE ERGATIVE SUBJECT

PAOLA BENEDINI

Why the ergative subject is easily identifiable? The aim of this article is to demonstrate that also in ergative languages it is possible to identify that particular NP, which is the agent and, probably, indeed very often, the topic of high transitive clause. This NP is clearly signalled by the ergative case-marking and corresponds to the traditional nominative agent-subject regarding its semantic function. This hypothesis is possible only if you consider Dixon's definition of syntactic ergativity not satisfactory for some aspects.

A satisfactory definition of the term subject is undoubtedly problematic. Since a predicate has 'per definitionem' at least one argument this can be considered its subject, but only in a formal sense. So the relationship subject-predicate has to be considered only a logical, not a grammatical one and probably any other use of the term is misleading in some sense. My aim in this article is clearly restricted. I will try to define only a particular kind of subject (the term is here used tentatively only to refer to the logical argument of predication) which seems to be 'easily identifiable' not only in morphological ergative languages but in every language whenever it uses a highly transitive construction i.e. an ergative one. For this hypothesis, some preliminary suppositions are necessary:
a) Dixon's definition of syntactic ergativity is not satisfactory for some aspects
b) a prototypical ergative construction may be analyzed as being a highly transitive one
c) the agent of an ergative clause has to be more topical than its patient.

1. The definition of syntactic ergativity

Dixon's definition of syntactit ergativity is neither functional nor satisfactory for at least two reasons. There is not a real typological opposition between the ergative and the nominative syntax; and in Dyirbal language there are grammatical counterexamples that consider the ergative NP relevant for syntactic transformations. After Dixon's description of Dyirbal the interest in ergativity shifted from its earlier purely morphological focus towards a more comprehensive syntactic view. Thus Anderson divided all ergative languages in two types: 'surface ergative languages' where ergativity is purely morphological but syntactic rules continue to abide by the nominative notion of subject, and 'deep ergative languages' where the syntactic behaviour of the language is also governed by the categories of ergative or absolutive. (For example Dixon identifies the absolutive NP in Dyirbal

as the 'topic'). But it is important to underline that the nature of morphological ergativity is not as trivial as Anderson seems to affirm. It may be that some languages clearly cannot be characterized at the syntactic level in terms of ergative accusative continuum. It is difficult to be sure on this point: although many languages have been adequately dealt with at the level of morphology few have been properly described at the syntactic level. So the traditional opposition between ergative type languages and nominative/accusative type languages, based on syntax, is problematic and not exhaustive.

A language is said to show ergative characteristics if intransitive subjects (S) are treated in the same manner as transitive objects (O) and differently from transitive subjects (A). A typical example of an ergative case-system is Dyirbal (Dixon 1972); here A is marked by ergative case inflection as opposed to S and O which are marked by absolutive case. According to Dixon, language can also be said to have 'ergative syntax': some rules of coordination and subordination will treat O and S in the same way and A differently. So, some Dyirbal's transformations are designed to produce absolutive chains i.e. strings of clauses with coreferential surface absolutive NP. It seems that Dyirbal's speakers employ a syntactic mechanism, notably the ŋ *ay* rule, to create just such absolutive chains, when the underlying forms of the clauses do not have a shared absolutive NP. For example two clauses can be coordinated if they involve a common NP which is in surface S or O function in each clause. The occurence of the common NP in the second clause is usually deleted and the whole biclausal construction can involve one intonation group. But,if the syntactic condition about coordination is not met an antipassive transformation must be applied deriving a construction in which an underlying A NP is in derived S function to satisfy the coreferentiality condition on coordination. An antipassive transformation derives an intransitive sentence from an underlying transitive construction; the original A NP becomes S, the intransitive subject, the original O NP now takes dative case and the verb is marked by the antipassive derivational affix ŋ *ay*. Moreover, Dixon's thesis affirms that Dyirbal is particular in that all major transformations (relativization as well as complementation) treat S and O as if they were a syntactic unit. He also states that Dyirbal is undoubtedly more ergative at the syntactic level than at the morphological one. The absolutive NP has to be included in each sentence, but an ergative NP may be deleted. So in Dixon's terms, the ergative case is always marked in an absolutive system: likewise absolutive has a syntactic-like status which is very similar to the one nominative has in Latin: i.e. the case which includes S function is more often the unmarked term in the system and has zero realization.

I find this kind of definition at least incomplete and probably inaccurate.The marked /unmarked parameter for case identification is not functional enough in understanding the ergative clause construction. Probably it would be better to consider the fact that there are a number of ergative languages in which subjects that are also true agents receive a special marker as opposed to subjects that are not true agents i.e. intransitive subjects. This seems to be the other side of the aspect of Dixon's definition of ergativity. So it is certainly important that S and O are morphologically identified, but it is much more relevant and to be underlined that the true A is signalled by a different case marking, that is the ergative one. Probably the ergative subject máy be considered as a marked construction, but this does not mean that the particular NP here involved is not a subject. Perhaps this is the only kind of subject easily identifiable. However it is important to keep in mind that

the notion of true agent can be defined only in the context of a transitive clause not simply by referring to a single participant.

Moreover, analyzing Dixon's grammar of Dyirbal, Heath 1979 tries to show that in formal terms there is not substantial evidence for either ergative or accusative syntax being present. So in formulating the structural description and the structural change for each rule in Dyirbal syntax it seems that transitive objects and intransitive subjects do not constitute a unit in opposition to transitive subjects A; but on the other hand, there is little evidence that IS and A (the traditional subject) form a syntactic unit in the presence of transformations. In contrast to Dixon's claim that IS and P cannot be deleted in isolated sentence while A can, Heath affirms that there surely must be one overt nuclear NP in each simple clause: either agent or intransitive subject or patient.

Another crucial point concerns Dixon's hypothesis that the fundamental unit of discourse is the absolutive chain (topic chain), meaning a string of consecutive clauses sharing a common absolutive NP. According to Heath, to take this as the basic discourse unit is arbitrary. For example it is not always true that EQUI NP deletion applies only in absolutive chains: moreover some of the crucial transformations do not alter underlying case categories as neither do all the configurations which should have been converted into absolutive chains. So it seems more profitable to speak of coreference chains approximately as sequences of clauses lacking internal boundaries so that each pair of adjacent clauses has a shared NP in any nuclear case (IS, P, A)[1].

---

1. There is no overt coordinating particle in Dyirbal, such as the English *and*. Thus if we wish to conjoin 1 and 2 we find that the syntactic condition on coordination is not met:
   1. ŋuma banaga + n y u   'Father returned'
   2. yabu ŋuma + ŋgu bura + n   'Father saw mother'
   The antipassive version of 2 is:
   3. ŋuma bural + ŋa + nyuyabu + gu 'Father saw mother'
   Now 1 and 3 can be coordinated in either order (Dixon 1972).
   According to Heath (1979) we need other rules to understand Dyirbal syntax: False Reflexive Rule: delete underlying P and consequently reshape the rest of the clause; the verb is intransivized by adding -Refl, while A becomes surface IS. Now this rule can be combined with Indefinite NP Deletion so that in a simple sentence A or P are deletable and IS is never so.

   4. bayi ja ŋ gay-mari-ñu   bagum +wujungu
      he-IS   eat-Refl-Pres   the   fruit-Dat

   P is a 'demoted' NP and takes dative case, it can be deleted if indefinite or unimportant:

   5. bayi   ja ŋ gay-mari-ñu
      he-IS   eat-Refl-Pres

   Heath's definition of ŋ ay rule:
   Let S_m be an underlying transitive clause preceded by a clause S_{m-1} in the same coreference chain. If A_m (i.e. the A np of S_m) is coreferential to any nuclear NP in S_{m-1}, then S_m undergoes ŋ ay antipassive. So in 6 the rule does not apply:

In short, Heath's hypothesis criticizes the typological classification of Dyirbal as an ergative language since it is impossible to find a single inequivocal example where a transformational process treats IS and P as a unit and excludes A. Moreover Dixon's claim that Dyirbal syntax can be understood in the light of the coreferential absolutive targets provides no insight into such important rules as EQUI and Indefinite NP deletion and so on. On the other hand, Heath does not think that Dyirbal may be syntactically accusative. His approach to grammar consists in considering grammatical systems as functionally integrated mechanisms adapted to achieve certain goals i.e. to maintain referential clarity. So the tendency of languages with ergative morphology to develop antipassives and use them for various purposes seems to be due to the fact that antipassive presents the sharpest formal contrast with unaltered transitive sentences. An antipassive rule would be more strongly marked than a passive one and its function would be clause-internal permitting the omission of an underlying P which is not needed. The point is that the development and functional extension of an antipassive rule need not be construed as an analogical pattern modelled on a morphological arrangement. So it seems that the definition of syntactic ergativity is more naturally and immediately linked to the parameters of transitivity and agentivity than to those of antipassive transformation and absolutive chains.

2. Transitivity and Ergativity

Let us suppose we redefine the ergative case-marking in functional terms; in this language type the morphology is certainly sensitive to the transitivity of the clause. Transitivity is really a fundamental parameter in grammar.

In every language the opposition transitive/intransitive is the one taken for granted

---

6. ŋaya bayi   + yara   balga-n   ⧣   walmbi-n
   I-A   the   man-P   hit-Past   lift-Past
   'I hit the man and (I) lifted him up'

Although the A of the second clause is coreferential to the A of the first clause, the P NP are also coreferential and so the rule is blocked.

A coreferential surface A NP in adjoining clauses in a coreferential chain regularly undergoes EQUI NP deletion:

7. ŋaya bala + yugu   yuba-n ⧣ balan + yugumbil   iilwa-n
   I-A   the   stick P   put down   the   woman P   kick Past
   'I put down the stick and (I) kicked the woman'

8. bala   + yugu   baŋgul   + yaraŋgy nudi-n   bayi   ñalŋga   bunju-n
   the   stick P   the   man A   cut Past   the   child P   spank Past

'The man cut the tree and (the man) spanked the child'

Because a P NP occurs in the second clause in each case, the deleted NP must be the A of the second clause. To recover the referential network we must decide whether the A or P of the first clause is the coreferential NP. Since the ηay rule has not been applied in these examples, the only other possibility is that $A_1 = A_2$.

in defininig ergativity itself: the division between transitive clause which has a patient-like direct object and displays ergative morphology and intransitive clause which has no such object and do not display ergative morphology. So the core of the semantic phenomenon of transitivity may be summarized by three major properties of the clause, pertaining separately to the agent, patient and verb:

Agent: a prototypical transitive clause has a visible, salient, intentional and controlling agent-cause

Patient: a prototypical transitive clause has a visible, salient, non-intentional, non-controlling patient result

Verb: a prototypical transitive clause has a compact, perfective, realis verb.
What one finds in split ergative languages is that the ergative morphology is sensitive, at some point, to the degree or strength of each one of these semantic properties of the clause.

A very similar approach to the definition of transitivity is due to Hopper and Thompson (1980). They try to define in universal terms the intuitive understanding of the notion of transitivity i.e. a global property of an entire clause, so that an activity is transferred from an agent to a patient. They identify the following parameters of transitivity each of which suggests a scale according to which clauses can be ranked:

1. Participants: no transfer at all can take place unless at least two participants are involved.

2. Kinesis: actions can be transferred from one participant to another; states cannot.

3. Aspect: an action seen from its endpoint, i.e. telic action, is more effectively transferred to a patient than one which is atelic. The activity has to be totally completed.

4. Punctuality: actions carried out with no obvious transitional phase between inception and completion have a more marked effect on their patients than actions that are inherently continuous.

5. Volitionality: the effect on the patient is typically more apparent when the agent is presented as acting purposefully.

6. Affirmation: this is the affirmative/negative parameter.

7. Mode: this refers to the distinction between realis and irrealis encoding of events. An action that directly corresponds to a real event is more effective than one that does not.

8. Agency: it is obvious that participants that are high in agency can perform a transfer of an action in a way that ones low in agency cannot.
Referring to objects:

9. Affectedness of O: the degree to which an action is transferred to a patient is a function of how completely the patient is affected.

10. Individuation of O: refers to the distinctness of the patient from the agent and from its own background.
An action can be more effectively transferred to a patient which is more individuated than to one which is less so; thus a definite O is often viewed as being more completely affected than an indefinite O.

Now the presence in a clause of all these parameters constitutes the highest possible degree of transitivity, so we can define every sentence in every language as more or less

transitive using these properties. This is a particular view of transitivity; it is a relationship which involves the whole clause and is not restricted to one constituent or pair of constituents. Consequently the presence of an overt O is only one feature of a transitive clause; it coexists with other defining properties such as agency, kinesis, volitionality etc. and just as a clause may have an overt second participant and still be aligned with the intransitive clause so it may also lack a second participant and yet have transitivity features. For (this is very interesting in Hopper and Thompson's hypothesis) transitivity is not dichotomous, but it is a continuum; so it then follows that clauses lacking an overt O must be settled somewhere on this continuum, but it does not necessarily follow that such a clause is situated at the extreme intransitive end.

A rather typical situation in ergative languages is that the canonical ergative clause signals several or all the high transitive features, while the antipassive clause signals one or more of the low transitivity features. Characteristically, Hopper and Thompson find these contrasts between ERGATIVE and ANTIPASSIVE:

verb codes two participants	verb codes only one participant
perfective aspect	imperfective aspect
total involvement of O	partitive O
definite O	indefinite O
kinetic/volition verb	stative/involuntary verb
active participation of A	passive participation of A

Moreover according to Anderson 1976 in some ergative languages there are numerous pairs of ergative/antipassive and they differ systematically in the following way: the antipassive form in each case indicates that the action is carried out less completely, less successfully, less conclusively etc., and that the object is less completely, less directly, less permanently etc. affected by the action.

So the global impression is that the ergative construction has the 'hall marks' of high transitivity, the antipassive those of low transitivity. In the ergative type the action is more intense, the involvement of the A is more deliberate; the O is specified and more completely affected. Another important parameter is the aspect that is systematically correlated with the degree of transitivity of the verb: if the aspect is perfective the interpretation - other things being equal - has properties allowing the clause to be classified as more transitive. In a considerable number of languages an ergative construction is limited to perfective or praeterit environments, while a non ergative type is used in imperfective environments (for example in Hindi and Georgian).

The transitivity of the clause is reduced when there is an anomalous A-O relationship: when the O is higher than A in animacy hierarchy. This claim is not a contradiction of the other claim that a clause is more highly transitive if its A is more highly agentive or its O more highly individuated. These statements are in fact mutually supporting since it is the reduction of the agency of the A which accounts for the anomaly in the A-O relationship, not the fact that the O is high in hierarchy. It is interesting to note that the proposals of Comrie and of Hopper and Thompson one make opposite predictions with respect to the marking of 'subject' and yet empirical facts exist to support both positions. Comrie's

claim implies that some languages will mark subject NPs just when they are inanimate/indefinite since this is the atypical situation for subjects: for example a number of Australian languages have a special marker for subjects low in animacy. On the contrary, Hopper and Thompson's hypothesis maintains that some languages will mark subjects NP only when they are true A, in a way parallel to how O NP are marked just when they are true O.

So there are a number of languages of ergative type in which subjects which are also A receive a special marker as opposed to subjects which are not true A. So the ergative case permits the recovery of semantic-syntactic function of that particular NP and also automatically predicts some other grammatical properties of that clause according to transitivity features. The notion of true agent i.e. the ergative subject can be defined only in the context of a transitive clause and not in referring simply to a single participant. Probably the 'ergative subject' can be considered as a prototypical multifactorial notion not only in ergative languages, but indeed in every language where a clause presents this kind of transitivity properties.

To sum up, it seems better to define syntactic ergativity as a high transitive construction with a particular kind of subject which is not only a clear volitional agent, but it has many other properties strictly linked to the degree of transitivity of its clause.

3. The agent-topic of an ergative clause

An interesting problem is represented by a group of languages that seem to consider the patient more topical in discourse than the agent. Probably this kind of affirmation is a consequence of Dixon's idea of ergative syntax. After Schachter's analysis of Tagalog it has been observed that in Tagalog goal-topic construction, the goal-patient retains several properties usually reserved for traditional subject in other languages. This same sort of observation has been made with regard to patients in transitive clauses of ergative languages and moreover this have been compared to passive constructions in clearly non ergative languages because of the subject-like properties which the patient seems to display. This is disturbing for the identification of the 'ergative subject' which naturally would have to be the agent and probably, indeed very often, the topic. But, according to Cooreman, Fox and Givon's hypothesis two kinds of text-based measurements have proved useful in evaluating the degree of topicality (of topic importance) of particular NP in discourse:

1. *Referential Distance:* here one measures the gap in a number of clauses between the present occurrence and the last preceeding occurrence of the topic. More continuous, important or topical participants will exhibit on the average smaller referential distance values (highest topic—continuity value = 1). This is a measure of anaphoric continuity-/predictability of the topic.

2. *Topic Persistence:* here one measures the number of contiguous subsequent clauses in which the participant NP remains a semantic argument, following the present occurence. More continuous, important, topical participant will exhibit on the average larger topic persistence value with the lowest topic continuity value, being zero by definition.

So, on this basis, Cooreman et al. intend to demonstrate that many discourse based measurements support the conclusion that Tagalog is substantially an ergative language

and they also show that (if topicality is a discourse -based notion) Tagalog is a surface er-
gative language with the agent of the ergative clause clearly more topical than its patient.
Futhermore, they intend to evaluate the degree to which Tagalog may still be considered
'syntactically' ergative in spite of the ergative agent NP. The text-count results are that
nearly 60% of transitive events are coded in discourse by the ergative construction, with
the rest spread amongst the other constructions. Moreover the agent in the ergative is co-
ded 75% by ∅ anaphora or pronouns while the patient is on average coded only 33% by
∅ anaphora or pronouns. The Tagalog passive displays a dramatic shift in these frequen-
cies, to 55% for agents and 54% for patients. By this rough measure of the topicality of
NP the ergative construction in Tagalog clearly identifies itself as one where the agent is
much higher in topicality than the patient, while in the passive construction the topicality
of the arguments is roughly equal.The construction tentatively labeled passive thus dis-
plays one of the more conspicuous and universal properties of passive, namely a low re-
ferentiality of the agent and a high ration of agent deletion, while only 15% of the agents
in the ergative construction are deleted, a full 53% of the agents in the passive are deleted.
These results are corroborated by those of the referential distance and topical persistence
parameters.

In the ergative construction the agent persists with more than twice the frequency
than the patient and is thus once again shown to be much more topical than the patient.
So, according to this definition of ergativity, in contrast to that of Dixon's, Tagalog is an
ergative language and indeed the ergative clause type is the most common one used in co-
ding semantically transitive events in discourse. If the topicality status of the NP ar-
guments in the various clause type is defined in a discourse based manner, then Tagalog is
a surface 'morphological' ergative language. The problem lies in Dixon's definition of er-
gativity which considers the patient absolutive as topic and as more relevant in syntactic
operations: that is, for example, coreferential absolutive chains. Cooreman et al. interpret
the relatively few instances where the coreference in the second clause pertains to the pa-
tient NP of the first clause, as a topic switch, while they see the overwhelming majority of
cases (75%) as representing instances where the agent persists as the clausal topic.

This solution is reached irrispective of the typological nature of Tagalog which is
probably an active-type language. But, this hypothesis represents a very alternative ap-
proach to the problem of defining ergativity, considering transitivity as strictly linked to it
from a pragmatic point of view.

4. Conclusions

Considering ergative languages from a point of view other than Dixon's definition, it
seems clear that the ergative case-marking may have an important function in signalling
one particular subject. So the morphologically 'ergative' or 'nominative' subject has to be
the agent and this agent has to be more topical than its patient in the transitive con-
struction of every type of language. The problem is rather about the 'absolutive' subject,
which is lower than ergative in agency and is not identified clearly by its case-marking.

So, if we assume that ergativity is strictly linked to the transitivity features (if transi-
tivity is a continuum) we can imagine a prototypical clause structure, probably an univer-
sal one,where the first argument is the intentional controlling agent causer, the verb is

highly transitive, perfective,realis and punctual, and the second argument is an individuated, distinct, completely affected, non intentional patient result. So, if these conditions are met (considering the ergative morphology only as a clearer signal of the NP semantic properties), the prototypical clause may be universally defined as highly transitive i.e. ergative and its first argument as the ergative subject.

Paola Benedini
Università Degli Studi di Bologna
Istituto di Glottologia
Via Zamboni 16
40126 Bologna

BIBLIOGRAPHY

Anderson, S. 1972: «On the syntax of ergative languages». In Heilmann (ed.), *Proceedings of the Eleventh International Congress of Linguists,* Bologna: Il Mulino, 73-77.
Anderson, S. 1976: «The notion of subject in ergative languages». In Li (ed.) 1976, 1-23.
Anderson, J. 1979: *On being without subject.* Bloomington, Indiana: I.U.L.C.
Blake, B.J. 1976: «On ergativity and the notion of subject». *Lingua* 39, 281-300.
Chafe, W. 1976: «Giveness, contrastiveness, definitess, subjects and topic». In Li (ed.) 1976, 25-55.
Comrie, B. 1972: «The ergative: Variation on a theme». *Lingua* 39, 239-253.
Comrie, B. 1978: «Ergativity». In W.P. Leheman (ed.) *Syntactic typology: Studies in the phenomenology of language,* Austin: University of Texas Press, 329-394.
Comrie, B. 1981: *Language universals and linguistic typology.* Oxford: Blackwell.
Cooreman, A. — Fox, B. — Givón, T. 1984: «The discourse definition of ergativity». *Studies in Language* 8, 1-34.
Cruse, D.A. 1973: «Some thoughts on agentivity». *Journal of Linguistics* 9, 11-23.
DeLancey, S. 1981: «An interpretation of split ergativity and related patterns». *Language* 57, 626-657.
DeLancey, S. 1984: «Notes on agentivity and causation». *Studies in Language* 8, 181-213.
Dixon, W. 1972: *The Dyirbal language of North Queensland.* Cambridge: Cambridge University Press.
Dixon, W. 1979a: «Ergativity». *Language* 55, 59-138.
Dixon, W. 1979b: «Corrections and comments concerning Heath's 'Is Dyirbal ergative?'». *Linguistics* 17, 1003-1015.
Foley, W. — Van Valin, R. 1984: *Functional syntax and universal grammar.* Cambridge: Cambridge University Press.
Heath, J. 1979: «Is Dyirbal ergative?». *Linguistics* 17, 401-463.
Heath, J. 1980: «Dyirbal ergativity: counter-rejoinder to Dixon». *Linguistics* 18, 505-521.

Hopper, P.J. 1979: «Aspect and foreground in discourse». In T. Givón (ed.), *Discourse and Syntax*. Syntax and Semantics 12, 213-241.

Hopper, P.J. — Thompson, S. 1980: «Transitivity in grammar and discourse». *Language* 56, 251-299.

Keenan, E.L. 1976: «Towards a universal definition of 'subject'». In C. Li (ed.) 1976, 303-333.

Klimov, G.A. 1983: «On contentive typology». *Lingua e Stile* 18, 327-341.

Li, C.N. 1976: *Subject and topic*. New York: Academic Press.

Moravcsik, E.A. 1978: «On the distribution of ergative and accusative patterns». *Lingua* 47, 233-279.

Payne, T.E. 1982: «Role and reference related subject properties and ergativity in Yup'ik Eskimo and Tagalog». *Studies in Language* 6, 75-106.

Plank, F. (ed.) 1979: *Ergativity: Towards a theory of grammatical relations*. New York: Academic Press.

Schachter, P. 1976: «The subject in Philippine languages». In C. Li (ed.) 1976, 491-518.

Taylor, A. 1976: «'Ergative-based' or 'Transitive-based'?». *Foundations of Language* 14, 1-17.