

Linguistic Typology: Valency Changing Operations

A short revision: valency, semantic roles, subject assignment

Verbal valency is about how many arguments a verb requires. Most verbs have between one and three arguments. These arguments can fulfil a variety of semantic functions, also known as semantic roles. Since there are as many semantic roles as there are verbs, and since it does not make sense to mark all these hundreds of roles in different ways, every language has a small number of macro-roles. Each macro-role comprises several semantic roles. Among the macro-roles there are agent, patient, recipient, beneficiary, etc.

The subject is that constituent which is the vantage point from which an event is looked at. Not every language has subjects and objects. Active-inactive languages seem to distinguish only between semantic macro-roles. Most other languages, however, need one such vantage point for each event.

Subject assignment follows semantic principles. If a verb takes only one constituent, that will be the subject. If it takes more than one constituent, the most agentive one will be the subject in the unmarked voice.

Valency changing operations fall into two categories: those which reduce the number of arguments and / or change the subject are called voice alternations; those which increase the number of arguments with or without changing the subject do not have a general cover term, but are called causatives, applicatives, etc., depending on their semantic functions.

Active and passive in English; promotion and demotion

English has two voices, the active and the passive:

- (1) *John ate the apples.*
- (2) *The apples were eaten (by John).*

In Ex. 1, *John* is the subject and *the apples* is the object. In Ex. 2, *the apples* is the subject and *John* is a non-obligatory constituent. But despite these differences, the sentences also have much in common: in both of them, *John* is the agent and *the apples* is the patient. Such pairs are equivalent concerning their truth conditions. For this reason, J. Wackernagel, *Vorlesungen über Syntax*, vol. 1 (1926: 135) famously stated that the passive is an unnecessary luxury some languages have. But if the passive is just an unnecessary luxury, why do so many languages have it?

The answer must be that active and passive differ pragmatically. That also explains why not all active sentences can have passive counterparts:

- (3) *John is too ill to work, but his brother supports him.*
- (4) *John is too ill to work, but he is supported by his brother.*
- (5) *John supports the Liberal Democrats.*
- (6) *?The Liberal Democrats are supported by John.*

Exx. 3 and 4 show that *support* can normally be used in both active and passive. Ex. 5 also contains the verb *support*, but the passive sounds odd (Ex. 6). Why? Subject assignment is about perspective. The subject is that constituent from whose perspective an event is looked at. In Exx. 3 and 4, John and his brother are equals and it does not matter from whose perspective you look at the event. But Ex. 5 is different. Unless John is a millionaire, his support is of minor significance for the party. You can view the event from his perspective: he is giving a certain percentage of his money to the party, so he is supporting it. For this it does not matter if this party is very small or very big. But it is difficult to view this event from the party's perspective and still regard it as real support. For the party, John is just one of thousands of supporters, and unless his contribution is substantial the passive with the party as subject is odd.

Let us look at Exx. 1 and 2 again. In the English passive, the former object becomes the new subject (promotion) and the former subject is either left out or becomes a peripheral constituent expressed in a *by*-phrase (demotion). This is in fact what happens in proto-typical passives, but there are also passives which are less prototypical:

- (7) *Ta'wach sivaatuch-i pakha-puga.*
man.SUBJ goat-OBJ kill.REMPAST
'The man killed the goat.'
- (8) *Sivaatuch-i pakha-ta-puga.*
goat-OBJ kill-PASS-REMPAST
'The goat was killed.'
- (9) *On danse.*
one dances
'There is dancing.'

Exx. 7 and 8 are from Ute (Uto-Aztecan) (see T. Givón (2001), *Syntax: An Introduction*, vol. 2, Amsterdam and Philadelphia, p. 130). Ex. 7 is active, Ex. 8 is passive. In the passive, the subject is removed, but the object is not promoted. Ex. 9 is from French. The verb is intransitive. In the impersonal construction, the subject is replaced by a general pronoun, but formally the construction remains active.

Keenan's generalizations concerning the passive

Here I shall present the most central claims made by Keenan 1985 (see reading list). Where I make claims about Latin, I refer to data presented in de Melo (forthcoming), *Zur Sprache der republikanischen carmina Latina epigraphica: Satzumfang, Satzkomplexität und Diathesenwahl* (Festschrift for Krummrey, ed. by P. Kruschwitz, p. 111–33).

If a language has passives with agent phrases, it also has passives without agent phrases.

Passives without agent phrases are more basic and thus more common than those with agent phrases. In English, German, French, Italian, and Latin both types of passives occur, but passives without agent phrases are more common. In Latin comedy (colloquial register), only 2.91% of the passives have agent phrases. In didactic poetry (formal register), 15.87% of the passives have agent phrases.

The tendency to leave out agent phrases can be observed in all registers, but is more marked in the colloquial ones. Why?

Note also that the passive in general is rarer in colloquial registers; of all passivizable verb forms, 9.80% are in the passive in comedy, but 26.25% are in the passive in didactic poetry. Similar tendencies can be observed in English, where frequency counts show that the passive is rare in spontaneous narratives (1%) and more frequent in academic prose (18%). Why?

Colloquial language has less complex sentences than formal language. In comedy, I found that 68% of all clauses are main clauses and 32% are subordinate. In formal didactic poetry I found 41% of main clauses and 59% of subordinate ones. This has consequences for the use of passives.

Comedy: the passive is mainly used to demote an unimportant or even unknown agent. Promotion of the agent is of secondary importance.

Formal poetry: the passive can also be used to demote an unimportant agent. But because the sentences are more complex, speakers / writers also try to keep the same subject throughout main and subordinate clauses as a way of maintaining clarity. One strategy to achieve identity of subjects is to passivize wherever promotion of an object to subject status has this effect (*The book fell off the table after Mark moved it to the edge* → *The book fell off the table after it was moved to the edge by Mark*). This is the main reason why formal texts with complex sentences have more passives across languages. Note also that the agent is simply demoted because that is a consequence of the promotion of the patient, not because the agent is unimportant or unknown. That is why formal texts have more passives with agent phrases.

If a language can passivize stative verbs, it can also passivize activity verbs.

The reason is that actives normally focus on the agent, while passives focus on the patient. Focus on the agent is compatible with activities and states, while focus on the patient normally implies a state. If a verb is already stative in the active, passive transformations are more difficult.

In English, both types can be passivized, but stative verbs (*like, know* etc.) can often be passivized only under special pragmatic conditions.

If a language can passivize intransitive verbs, it can also passivize transitive ones.

The passive transformation leads to the loss of the the original vantage point (subject); if the verb is transitive it is easy to assign a new vantage point, but if the verb is intransitive there is no new vantage point at all.

English cannot passivize intransitive verbs; Latin and German can, but the construction is rare:

- (10) *Es wurde getanz.*
it be.PAST-3SG dance.PARTICIPLE
'There was dancing.'

This construction is used if you want to leave the agent unspecified.

If a language has one or more passive, at least one can cover the perfective range of meaning.

Statistically speaking, past and perfective go hand in hand with focus on the patient. This is why perfective and past passives are more basic than imperfective and non-past ones. In a text count, I found that in Latin comedy the passive is four times more frequent among the past tenses than among present and future.

If a language has two or more basic passives, they are likely to differ semantically with respect to the aspect ranges they cover.

- (11) *The cup was broken.* = state
(12) *The cup got broken.* = activity

The subject of a passive verb phrase is never understood to be less affected by the action than when it is presented as the object of an active transitive verb.

This is because the subject is the vantage point from which an action is presented. Compare:

- (13) *John was watching the children while they were playing in the streets.* = simply looking at them, they are not affected
(14) *While they were playing in the streets, the children were being watched by John.* = either John is making sure that they are safe, or he has some unhealthy obsessions

This also explains the oddity of *The Liberal Democrats were supported by John*. The party is not affected a great deal.

Distinct passives in a language may vary according to the degree of affectedness of the subject and whether it is positively or negatively affected, though this variation seems less widely distributed than that of aspect.

see below on Chamorro

If a language can passivize transitive verb phrases taking sentential objects, it can passivize ones taking lexical noun phrase objects.

The new subject is the new vantage point from which you look at a state of affairs. It is easier to look at a state of affairs from the vantage point of a concrete entity than from the vantage point of a clause.

(15) *Everybody knows John.* → *John is known by / to everybody.*

(16) *Everybody knows that the earth is round.* → *That the earth is round is known by / to everybody.* / *The fact that the earth is round is known by / to everybody.*

Ignore the alternation between *by* and *to*. This has to do with the fact that *know* is a stative verb, and stative verbs are more difficult to passivize. In Ex. 16 the first passive transformation is accepted by most speakers, although some prefer the second one with *the fact that*. This second version subordinates the clause to a noun phrase, so that the passive subject is formally a noun phrase, not a clause.

If a language has basic passives, it always passivizes three-place verbs in such a way that the derived subject is the patient of the active verb. Passives in which the recipient is the subject may or may not exist.

Compare:

(17) *A book was given to John.*

(18) *John was given a book.*

Ex. 17 has an equivalent in German, Ex. 18 would be ungrammatical.

Actually, what can become the subject of an alternative voice is not random, but follows the semantic function hierarchy:

agent > patient > recipient > beneficiary > instrument > location > time

The agent is most unmarked as subject and is the regular subject in the active. Other semantic functions can become subjects in other voices, but languages differ as to where the cut-off point lies. Compare the following table, based on S. C. Dik (1997), *The Theory of Functional Grammar*, part 1: *The Structure of the Clause* (Berlin and New York), p. 267:

Language	Agent	Patient	Recipient	Beneficiary	Instrument	Location	Time
Dutch	yes	yes	no	no	no	no	no
English	yes	yes	yes	no	no	no	no
Sundanese	yes	yes	yes	yes	no	no	no
Maranao	yes	yes	yes	yes	yes	no	no
Malagasy	yes	yes	yes	yes	yes	yes	no
Cebuano	yes	yes	yes	yes	yes	yes	yes

Note: Dutch and English are West Germanic languages; the other languages are Austronesian.

Voice: the antipassive in Dyirbal

Dyirbal is a Pama-Nyungan language spoken in North Queensland, Australia. The most extensive description is R. M. W. Dixon (1972), *The Dyirbal Language of North Queensland* (Cambridge).

Morphologically, Dyirbal is a typical Australian language: common nouns mark the agent of a transitive verb (A) with the ergative case, whereas the object of a transitive verb (P) or the subject of an intransitive verb (S) is in the morphologically unmarked absolutive case; by way of contrast, pronouns in A or S function are in the morphologically unmarked nominative case, while P is in the accusative.¹ The following table illustrates these patterns:

	S	A	P
we (dual)	<i>ŋali</i>	<i>ŋali</i>	<i>ŋali-pa</i>
rainbow	<i>yamani</i>	<i>yamani-gu</i>	<i>yamani</i>

Table: Marking of core cases among nouns and pronouns

S and A have the same form for pronouns, while among the nouns S and P look the same. However, even though among the pronouns S and A differ from P morphologically, it is S and P that behave in similar ways syntactically, just as among the nouns. For this reason we can ignore the distinct morphological patterns of the pronouns in the following exposition.

Exx. 19 and 20 illustrate transitive and intransitive clauses:²

- (19) *Balan* *ɖugumbil* *baŋgul*
 CLASSII.ABS woman.ABS CLASSI.ERG
yaŋa-ŋgu *balgan*.
 man-ERG hit.NONFUT
 ‘The man is hitting the woman.’

¹ *Who* has three different forms for S, A, and P.

² Note that Dyirbal distinguishes future and non-future tenses; whether a non-future tense should be interpreted as past or present depends on the context.

- (20) *Bayi yaɾa ɲinaɲu diban-da.*
 CLASSI.ABS man.ABS sit.NONFUT stone-LOC
 ‘The man sat on the stone.’

Sentence 19 is transitive; it has an ergative A and an absolutive P. Sentence 20 is intransitive, so S is in the unmarked absolutive case.

Equi-NP deletion in English and Dyirbal is different because in English A and S are coreferential, while in Dyirbal P and S are coreferential. Compare:

- (21) *The man is now hitting the woman, but sat on the stone before.*

→ The man sat on the stone before.

- (22) *Balan ɟugumbil baɲgul*
 CLASSII.ABS woman.ABS CLASSI.ERG
yaɾa-ɲgu balgan; ɲinaɲu diban-da.
 man-ERG hit.NONFUT sit.NONFUT stone-LOC
 ‘The man is hitting the woman and (x) sat on the stone.’

→ x = the woman

How can you co-ordinate the two clauses and get the same meaning as in English? For this you need to transform the A of the transitive clause into S. This is what the antipassive does. Antipassives remove P or give P an oblique case form rather than absolutive case marking, and as a consequence the clause is now intransitive; A becomes S. Compare:

- (23) *Bayi yaɾa bagun*
 CLASSI.ABS man.ABS CLASSII.DAT
ɟugumbil-gu balgal-ɲaɲu.
 woman-DAT hit-NONFUT.ANTIPASS
 ‘The man is hitting the woman.’

Ex. 23 refers to the same state of affairs as Ex. 19, but is syntactically different. Ex. 23 is intransitive; the subject is in the absolutive case and the object is in an oblique case, the dative as here or the instrumental. Such a construction is generally referred to as an antipassive.

Now compare:

- (24) *Bayi yaɾa bani-ɲu.*
 CLASSI.ABS man.ABS come-NONFUT
 ‘The man came here.’

If we want to co-ordinate this clause with Ex. 19 above, the A in Ex. 19 has to become an S by antipassivization first. Ex. 23 is the antipassivized version of Ex. 19. We can now co-ordinate Exx. 24 and 23 and leave out the second S, as in Ex. 25:³

³Note that Dyirbal does not have a connective element like English *and*.

- (25) *Bayi yaɾa bani-ɲu*
 CLASSI.ABS man.ABS come-NONFUT
 ∅ *bagun ɟugumbil-gu balgal-ɲapu.*
 (S) CLASSII.DAT woman-DAT hit-NONFUT.ANTIPASS
 ‘The man came here and (he) hit the woman.’

According to Dixon’s account (1972: 66), antipassivization in Dyirbal cannot occur discourse-initially; its main function is to enable speakers to co-ordinate sentences while keeping the same topic (the S or P constituent).

The S / P (‘absolutive’) relation seems to be more central to the Dyirbal system than the S / A (‘subject’) relation. In the Dyirbal antipassive, the central P is demoted and marked by an oblique case (dative or instrumental), whereas the former A is promoted and becomes the central absolutive constituent (S).

A complex voice system: Chamorro

Chamorro is an Austronesian language spoken on the Mariana islands. In my account, which is mainly based on Cooreman (1987), I shall restrict myself to semantically transitive states of affairs, that is, those in which there is an agent and a patient.

We can distinguish between four different voices: the active (called ergative / irrealis by Cooreman), the MA-passive, the IN-passive, and the antipassive.⁴ The following sentences illustrate these voices:

- (26) *Ha-fahan si Juan i niyok.*
 ERG.3S-buy ART John ART coconut
 ‘John bought the coconut.’
- (27) *Ni-na’-ma’a’ñaο hao ni estoria.*
 PAS-CAUS-afraid A.2S OBL story
 ‘You were frightened by the story.’
- (28) *Todu na taotao ni mang-gaige*
 all LINK people REL PL-be
Guam guihi na tiempo man-ma-takpangi.
 Guam there LINK time PL-PAS-baptize
 ‘All the people who were in Guam at that time were baptized.’
- (29) *Man-offresi si Juan nu hagu ni salape’.*
 A.P.-offer ART John OBL EMPH.2S OBL money
 ‘John offered the money to you.’

Ex. 26 is a straightforward active sentence. The regular word order is VSO, so the verb is followed by the agent *si Juan* ‘John’ and the patient *i niyok* ‘the coconut’. Ex. 27 and 28 are both passives. Ex. 27 is the IN-passive;⁵ the

⁴The ergative and the irrealis employ different verb prefixes (Cooreman (1987: 35)), but differ mainly in temporal reference and modality rather than in argument structure.

⁵The prefix *ni-* is a variant of the infix *-in-*.

pronoun has been promoted to subject function, while the inanimate agent, *ni estoria* ‘the story’, has been demoted and is marked as oblique. Ex. 28 is the MA-passive; *todu na taotao* ‘all the people’ is the patient and has become the subject, while the agent is left unspecified. Ex. 29 is an antipassive. Both agent and patient are present in the sentence, but the patient has become a less central constituent and is marked as oblique.

What makes Cooreman’s text-based study so outstanding is the fact that it renders the choice between the different voices more or less predictable. This choice depends on the topicality of the constituents. Topicality could for our purposes be defined as importance for the discourse. Cooreman establishes a way of calculating the topicality of a constituent: she measures how many clauses intervene between the constituent itself and its last mention (‘look-back’) and also how many clauses come between the constituent itself and its next mention (‘look-ahead’); the shorter these distances are, the greater the topicality of a constituent.⁶

The following table, based on Cooreman (1987: 76), summarizes how topicality and voice are interrelated:

Voice	Relative topicality
Antipassive	agent >> patient
Active	agent > patient
IN-passive	agent < patient
MA-passive	agent << patient

Table: Voice and the topicality of the core constituents

From this table we can see that the active and the antipassive are chosen if the agent is more important than the patient, whereas one of the passives is chosen otherwise. The antipassive is preferred over the active if the patient is far less topical than the patient. The MA-passive is preferred over the IN-passive if the patient is far more topical than the agent.

Causatives

In a causative construction, the number of constituents increases by one — the causer. From a purely formal point of view, we can distinguish between analytic causatives, causatives with derivational morphology, and lexical causatives:

(30) *John laughed.* → *Mike made John laugh.* = analytical, periphrasis

(31) Turkish: *öl* ‘die’ → *öl-dür* ‘kill’ → *öl-dür-t* ‘cause to kill’ = derivational suffixes

⁶For a more detailed discussion of this method, which works well for third persons, cf. Cooreman (1987: 11-21). There are certain limitations; discourse participants for example are inherently topical, so the quantitative method cannot be applied to them.

(32) *John died.* → *Mike killed John.* = lexical, suppletive verb

From a semantic point of view, we can distinguish between direct causation and indirect causation:

(33) *John made Jack leave.* = direct

(34) *John let Jack leave.* = indirect, permissive

Correlation between form and meaning:

If a language has more than one causative, the analytic forms tend to be used for indirect causation, the lexical ones for direct causation, and the morphological ones for events in between.

Sometimes case marking for core constituents is also a means of distinguishing between direct and indirect causation; cf. Hungarian:

(35) *Én köhögtettem a gyerek-et.*
I caused-to-cough the child-ACC
'I made the child cough.'

(36) *Én köhögtettem a gyerek-kel.*
I caused-to-cough the child-INSTR
'I made the child cough.'

Accusative marking is used for fully affected objects and hence direct causation. Maybe the speaker in Ex. 35 clapped on the child's back so that he or she could not help coughing. Instrumental objects are not fully affected and hence indicate indirect causation. Perhaps the speaker of Ex. 36 persuaded the child to simulate.

What happens to the causee? If possible it becomes the direct object. Cf. Turkish:

(37) *Hasan öl-dü.*
Hasan die-PAST
'Hasan died.'

(38) *Ali Hasan-ı öl-dür-dü.*
Ali Hasan-ACC die-CAUS-PAST

In Ex. 37 the intransitive subject is in the nominative. In Ex. 38 the causer is in the nominative and the former subject becomes direct object.

What if there is already a direct object? English is unusual in that it allows to direct objects:

(39) *John made me* (causee, direct object) *hate him* (direct object).

In Turkish and many other languages, the causee must be in some oblique case if the direct object role is already occupied:

- (40) *Müdür mektub-u imzala-dı.*
 director letter-ACC sign-PAST
 ‘The director signed the letter.’
- (41) *Dişçi mektub-u müdür-e imzala-t-tı.*
 dentist letter-ACC director-DAT sign-CAUS-PAST
 ‘The dentist made the director sign the letter.’

Ex. 40 is a normal transitive clause with nominative subject and accusative object. In Ex. 40 there is a new causer in subject role, but the accusative object role has already been taken. The causee must be expressed in the dative. If that role is also taken, Turkish employs a postpositional phrase for the causee.

Applicatives

Applicatives turn a peripheral constituent or a prepositional object into a direct object. Compare English *wait for someone* (prepositional object) vs. *await someone* (direct object).

My examples come from two languages of the Kru family (part of the Niger-Congo family, spoken in Liberia and Ivory Coast): Grebo (Western Kru) and Bete (Eastern Kru). For Grebo see G. Innes (1966), *An Introduction to Grebo* (London). For Bete see R. Zogbo (2004), *Parlons bété. Une langue de Côte d’Ivoire* (Paris) and H. Koopman (1984), *The Syntax of Verbs. From Verb Movement Rules in the Kru Languages to Universal Grammar* (Dordrecht and Cinnaminson).

Let us begin with Grebo. In Grebo the regular order is SVO and constituents other than subject and object are regularly marked by the use of postpositions. Ex. 42 shows that locative adjuncts actually can occur without postpositions:

- (42) *Ne yi-da nɔ ne ke London.*
 I see-PREHEST.PAST him AFFIRM there London
 ‘I saw him in London.’

However, as we can see here, if there is no postposition, the use of deictic *ke* ‘there’ is obligatory. In our next example, there is a postposition to mark the locative:

- (43) *Ne yi-da nɔ ne (ke) kae ye.*
 I see-PREHEST.PAST him AFFIRM (there) house in-front-of
 ‘I saw him in front of the house.’

Because of the postposition, *ke* ‘there’ is not obligatory.

In Ex. 44, the verb has an applicative suffix:

- (44) *Ne du-di bla sũ ne Mɔle.*
 I pound-APPL rice pestle AFFIRM Monday
 ‘I pounded rice with a pestle on Monday.’

As there is an applicative suffix in this example, the noun *sũ* ‘pestle’ can easily be identified as the instrument, especially since its lexical meaning makes this interpretation likely as well. Because of the applicative suffix, the instrument becomes the direct object of the verb, so no instrumental postposition is necessary.

Now let us look at Bete. Again we are dealing with an SVO language, and again constituents other than subject and object tend to be marked by postpositions. First some data from Bete as spoken in the region of Daloa:

(45) *atv gli ,bha †gbë †mõ.*
 father be LOC village in
 ‘Father is in the village.’

(46) *gwie , la † sɔ †mlin† †wv, †yrvbõ †mõ.*
 children DEM two be equal years in
 ‘These two children are equal in age.’

In Ex. 45 the postposition indicates ‘containment within’, while in Ex. 46 it could be translated as ‘in terms of’, which is a metaphorical extension of the concept of containment.

We do not know how many adpositions Bete has, but there are some verbal affixes whose use makes adpositions redundant. *-ni*, the applicative suffix, is one of them. In Ex. 47 it has instrumental function:

(47) *li-ni, cεε*
 eat-APPL hand
 ‘eat with one’s hand’

Here, the verbal affix already marks the function of the noun as instrument, so no separate postposition is employed.

Notice, however, that in other dialects of Bete, an applicative suffix on the verb is not sufficient to code an instrument; a postposition is used as well, which is typologically unusual. Compare the following example from the dialect spoken in Gbadi:

(48) *dí-l̄ nēmē gbèlè nú kú*
 cut-APPL meat knife with PARTICLE
 ‘Cut the meat with a knife.’

Here, the verb has an applicative suffix, and the instrument is in addition followed by the postposition *nú*. Now compare the following from the same dialect:

(49) *wá yī ðùdù kīyī sū kú dí-l̄*
 they FUT-A house behind tree PARTICLE cut-APPL
 ‘They will cut the tree behind the house.’

(50) *wá ðliē-lē nyíkɔ̀*
 they talk-APPL someone
 ‘They talked to someone.’

In Exx. 49 and 50, the same applicative suffix is used as in Ex. 48. In Ex. 49, it is employed because there is a locative adjunct in the sentence; but this adjunct is nevertheless marked by a postposition. In Ex. 50, the additional argument has the function of goal. In this case, no adposition is necessary.