THEMATIC PROTO-PROPERTIES AND ARGUMENT ENCODING IN THE ACTIVE-STATIVE LANGUAGES

Peter M. Arkadiev
Institute of Slavic Studies, Russian Academy of Sciences, Moscow
alpgurev@gmail.com, peterarkadiev@yandex.ru

1. Preliminaries: Semantics of ‘Semantic Alignment’


- which semantic properties are relevant for argument encoding with monadic verbs cross-linguistically?
- how are these properties interrelated and to which extent are they universal?

Answers which I am aware of:

- agentivity, affectedness, stativity, and telicity are crucial for argument encoding;
- these properties are independent, but intercorrelated.

A major theoretical proposal in a universalist vein (Primus 1999: 99 — 100; following Dowty 1991; cf. also Kibrik 1997):

(1) the more Proto-Agent (resp. Proto-Patient) properties an intransitive predicate entails with respect to its sole argument, the more likely the latter is to be encoded similarly to the Agent (resp. Patient) of the transitive predicate.


(2) Proto-Agent properties:
   a. volitional involvement in an event or state
   b. sentience (and/or perception)
   c. causing an event or change of state in another participant
   d. movement (relative to the position of another participant)
Proto-Patient properties:

a. undergoes change of state  
b. incremental theme  
c. causally affected by another participant  
d. stationary relative to movement of another participant

Proto-properties are independent of each other (see discussion by Dowty 1991: 572—574) but significantly intercorrelated (cf. Hopper & Thompson 1980 for correlations between various semantic parameters of transitivity).

A further question:

\[\text{does (1) really work?}\]

Answer:

\[\text{not quite!}\]

The goal of this paper is to show that (1) may be regarded only as a universal statistical tendency, but not as a generalization valid for particular languages, where finer-grained distinctions are at work.

2. Case Studies

2.1. Loma: a genuine ‘active-stative’ language

Loma is a Mande language spoken in Liberia. Primary source: Rude 1983.

Two sets of pronouns:

<table>
<thead>
<tr>
<th></th>
<th>‘Active’</th>
<th>‘Stative’</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Singular</td>
<td>Plural</td>
</tr>
<tr>
<td>1st</td>
<td>gà</td>
<td>gá (excl.)</td>
</tr>
<tr>
<td></td>
<td>dá (incl.)</td>
<td></td>
</tr>
<tr>
<td>2nd</td>
<td>jà</td>
<td>wá</td>
</tr>
<tr>
<td>3rd</td>
<td>tóa</td>
<td>tá</td>
</tr>
</tbody>
</table>

Dyadic predicates: <θ<sub>Agent</sub>: Active; θ<sub>Patient</sub>: Stative>. Examples:

\[(4)\]

\[gá \quad \text{té} \quad \gamma u \quad zu\]

we.EXCL.ACT they.ST see DUR

‘We see them’

Monadic predicates: case marking is determined exclusively by Aktionsart:

\[(5)\]

a. <Pred<sub>Stative</sub>: Stative>; b. <Pred<sub>Dynamic</sub>: Active>.

Examples:
Stative predicate:

(6) \( \text{gé} \ \beta \text{ala} \ \beta \text{e} \)  
we.EXCL.ST big DUR  
‘We are big’

Dynamic predicates:

(7) \( \text{gá} \ \text{li} \ \text{zu} \)  
we.EXCL.ACT go DUR  
‘We are going’

(8) \( \text{tóa} \ \text{lo} \ \text{zo} \)  
he.ACT fall DUR  
‘He is falling’

Incidentally, ‘fall’ entails more Proto-Patient properties with regards to its sole argument than ‘big’:

(9) ‘fall’: <P: change of state>

(10) ‘big’: <no>

Thus Loma constitutes a clear counter-example to (1).

2.2. Georgian: Telicity


Case marking in the Aorist (with certain simplifications; the so called ‘inversive’ verbs are excluded):

Dyadic verbs: \(<\theta_{\text{Agent}}: \text{Ergative}; \theta_{\text{Patient}}: \text{Nominative}>\). Examples:

(11) \( \text{glex-ma da-tes-a simind-i} \)  
peasant-ERG PRV-sow-AOR.3SG.SB corn-NOM  
‘The peasant sowed corn’

(12) \( \text{bavšv-ma gada-q’ar-a kv-eb-i} \)  
child-ERG PRV-throw-AOR.3SG.SB stone-PL-NOM  

(Non-stative) monadic verbs fall into two major classes roughly characterizable as Atelic and Telic (cf. Van Valin 1990).

Examples:
Agentive Atelic:

(13) $k'a$c-ma $i$-lap’ar*(a)
man-ERG PRF-talk-AOR.3SG.SB

‘The man talked’

(14) $gogo$-eb-ma $i$-tamaš-es
girl-PL-ERG PRF-play-AOR.3PL.SB

‘The girls played’

Proto-properties: <A: volition; sentience; causing; (movement); P: no>

Non-agentive Atelic:

(15) $c'q'al$-ma $i$-$du'\gamma$-
water-ERG PRF-boil-3SG.SB

‘The water boiled’

Proto-properties: <A: no; P: affected; stationary>

Agentive Telic:

(16) $bavšv$-eb-i $da$-sx$\delta$-nen
child-PL-NOM PRV-sit.down-AOR.3PL.SB

‘The children sat down’

(17) $k'a$c-i $a$-$\mu\bar{r}$-d-a
man-NOM INCH-sing-INCH-AOR.3SG.SB

‘The man began to sing’

Proto-properties: <A: volition; sentience; causing; (movement); P: change-of-state>

Patientive Telic:

(18) $c'q'al$-i $ga$-tb-a
water-NOM PRV-warm.up-AOR.3SG.SB

‘The water became warm’

(19) $k'a$c-i $mo$-k’vd-a
man-NOM PRV-die-AOR.3SG.SB

‘The man died’

Proto-properties: <A: no; P: change-of-state; (incremental theme); affected; (stationary)>

Generalization (1) does not work, again.

Case-marking rule for Georgian monadic verbs:

(20) a. $<\theta_{p,\text{change-of-state}}$: Nom>; b. $<\text{elsewhere: Erg}>$
2.3. Bats and Tabassaran: Volitionality

Bats and Tabassaran are North-East-Caucasian languages. Primary sources of data: Holisky 1987 for Bats, Kibrik 1985 for Tabassaran.

**Bats.**

Dyadic verbs: \( \theta_{\text{Agent}}: \text{Erg}; \theta_{\text{Patient}}: \text{Nom} \). Examples:

(21) \[ k'nat-ev \quad bader \quad dah''-daph-diē \]
    boy-ERG   child(NOM)   PRV-undress-AOR.3SG

‘The boy undressed the child’

(22) \[ badr-ev \quad surat \quad qoc'jiē \]
    child-ERG   picture(NOM)   hang-AOR.3SG

‘The child hung the picture’

Monadic verbs: case marking variation with pronominal subjects. Some verbs take only Nominative subjects, some only Ergative subjects, and some allow both. Cf. example of ‘fluid-marking’:

(23) a. \[ as \quad wože \]   b. \[ so \quad wože \]
    I(ERG)   fell       I(NOM)   fell

‘I fell (on purpose)’   ‘I fell (by accident)’

Verbs which take Nominative subjects:

(24) \[ maicdar \quad \text{‘be hungry’}, \quad dah'' \quad \text{δordar} \quad \text{‘freeze’}, \quad qerl'ar \quad \text{‘be afraid’} \] etc.

Proto-properties: various, **but not** \(<A: \text{volitional; causing}>\)

Verbs which take only Ergative subjects:

(25) \[ dat'ar \quad \text{‘run’}, \quad daqar \quad \text{‘come’}, \quad dadar \quad \text{‘swear’}, \quad axar \quad \text{‘bark’}, \quad lap'c'ar \quad \text{‘play’} \] etc.

Proto-properties: \(<A: \text{volitional; causing}>\) and possibly other.

Verbs which take both Ergative and Nominative subjects:

(26) \[ dah'' \quad daxar \quad \text{‘get drunk’}, \quad \text{δopdalar} \quad \text{‘hide’}, \quad dožar \quad \text{‘fall’}, \quad h''edar \quad \text{‘be late’} \] etc.

Ergative marking clearly implies volitionality of the subject:

(27) \[ \text{δopdalar} + \text{Nom}: \quad \text{‘subject comes to be hidden not because of anything she herself does, but because something moves in front of her’}; \]

(28) \[ \text{δopdalar} + \text{Erg}: \quad \text{‘subject does something which results in her becoming hidden, e.g. moves behind a barrier’} \]

Some peculiar cases:
(29) *dah’* davar + Nom: ‘to die’; + Erg: ‘to die because of either doing something in order to die or doing nothing to prevent death’

(30) *ga=rek’adalar* + Erg: ‘to run’; + Nom: ‘to run unwillingly, e.g. because of a very steep way’

Case marking rule for Bats monadic verbs:

(31) a. $\theta_{<A: volition>}$: Erg; b. elsewhere: Nom

Tabassaran.

Two sets of verb agreement affixes:

<table>
<thead>
<tr>
<th>Agentive</th>
<th>Patientive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Singular</td>
</tr>
<tr>
<td>1st</td>
<td>-za</td>
</tr>
<tr>
<td>2nd</td>
<td>-va</td>
</tr>
</tbody>
</table>

Dyadic verbs: $\theta_{Agent}$: Agentive; $\theta_{Patient}$: Patientive. Example:

(32) *uzu* *uvu* *kurc̣un-*za-vu

I you(SG) beat-1SG.AG-2SG.PT

‘I have beaten you’

Monadic verbs, like in Bats, fall into three major classes:

Verbs which take only Agentive agreement:

(33) *daqun-*za ‘I lay down’, *kižun-*za ‘I began to cry’, *kušun-*za ‘I came’ etc.

Proto-properties: $<A: volition; cause>$ and maybe other.

Verbs which take only Patientive agreement:

(34) *kabqun-*zu ‘I drowned’, *karrun-*zu ‘I froze’, *ergṛa-*zu ‘I got tired’ etc.

Proto-Properties: various, **but not** $<A: volition>$.

Verbs which may take both Agentive and Patientive agreement:

(35) *kužun-*za ‘I remained (voluntarily)’ vs. *kužun-*zu ‘I remained (against my will)’; *aqun-*za ‘I fell (intentionally)’ vs. *aqun-*zu ‘I fell (by accident)’; *hilirqun-*zu ‘I shook (on purpose)’ vs. *hilirqun-*zu ‘I shook (involuntarily)’ etc.

Case marking rule for monadic verbs in Tabassaran:

(36) a. $\theta_{<A: volition>}$: Agentive; b. elsewhere: Patientive

Bats and Tabassaran seem to differ in the degree of **pragmatic determination** of the $<volition>$ entailment: in Tabassaran the presence or absence of this entailment is more conventionalized.
2.4. Central Pomo: Affectedness


Case marking of pronouns: Agentive, Patientive and Oblique cases.

Dyadic verbs: $\theta_{\text{Agent}}$: Agentive; $\theta_{\text{Patient}}$: Patientive. Examples:

(37) a. $\wedge$: mú:tu $\tilde{\eta}$:čadiw
I(AG) he(PT) chased.away

b. mu:l  $\tilde{\eta}$:čadiw
he(AG) I(PT) chased.away

‘I chased him away’  ‘He chased me away’

Monadic verbs fall into three major classes cutting across Stative/Dynamic distinction.

Dynamic verbs taking Agentive subjects:


Proto-Properties: <A: volitional, causing> and possibly other.

Stative verbs taking Agentive subjects:


Proto-properties: most commonly <none>.

Dynamic verbs taking Patientive subjects:

(40) $\circ$: ló:ya ‘I fell’, $\circ$: madáts’čiw ‘I slipped’, $\circ$: šk’hé:nada ‘I’m getting well’, $\circ$: qamá:źleč ‘I got angry’ etc.

Proto-properties: <P: affected; change-of-state>, often <A: sentient>

Stative verbs taking Patientive subjects:

(41) $\circ$: kasíla ‘I’m cold’, $\circ$: kíts’čiw ‘I’m scared’, $\circ$: ńh:ál ‘I’m in pain’, $\circ$: mká:t ‘I’m surprised’ etc.

Proto-properties: <P: affected>, often <A: sentient>

Verbs which may take both Agentive and Patientive subjects:

(42) $\wedge$: sma mtí:če ‘I went to bed’ vs. $\circ$: sma mtí:čka ‘I must have fallen asleep’; $\wedge$: k’lúc’lu:w ‘I coughed (intentionally)’ vs. $\circ$: k’lúc’lu:w ‘I coughed (involuntary)’

Case marking rule for Central Pomo monadic verbs:

(43) a. $\theta_{\text{A: volition}}$: Agentive; b. $\theta_{\text{P: affected}; (A: sentient)}$: Patientive;
  c. <elsewhere: Agentive>

Rule ordering: a > b > c
3. Discussion and conclusions

As the data from the languages surveyed in 2 show, Primus 1999’s generalization (1) does not hold: it is usually not the overall balance of Proto-Agentive and Proto-Patientive properties the predicate entails with respect of its sole argument which determines the its morphosyntactic encoding, but a specific property from the sets in (2) and (3), which may override all other Proto-properties, notwithstanding their cumulative ‘weight’.

E.g., in Bats, if the speaker construes his/her sentence in such a way that the verb entails <A: volition>, its argument is encoded as Proto-Agent, regardless of any Proto-Patientive properties the same verb may entail (e.g. die or fall). Similarly, in Georgian, if the verb entails <P: change-of-state>, its argument is construed as Proto-Patient, even if it is genuinely agentive (e.g. stand up or begin to sing).

Moreover, the properties crucial for argument encoding may come from both sets of features, cf. an entailment constellation <A: sentient; P: affected> relevant for stative verbs in Central Pomo: it is precisely those predicates which entail these two properties for their argument which prototypically take Patientive subjects in Central Pomo.

Therefore, it is possible to argue that an adequate theory of ‘semantic alignment’ must incorporate not only a general principle like (1), but also a parametrization of it, i.e. it must allow individual languages to select from the universal set of Proto-properties a single specific semantic feature or feature constellation which will determine argument encoding.

Principles like (1) are useful in that they capture the basic intuition about the nature of ‘semantic alignment’: we do not find languages where a verb like ‘die’ encodes its argument as Proto-Agent and a verb like ‘work’ as a Proto-Patient. However, application of typological generalizations to individual languages is not always straightforward, because languages tend to give prominence to certain semantic features at the expense of other.

Such an approach allows us also to account for a well documented fact (cf. e.g. Merlan 1985) that in languages with ‘semantic alignment’ verb classes are often unequal in size and productivity: when some particular Proto-property is relevant for argument encoding, only those verbs which entail this property will take subjects marked as Agents (resp. Patients), whereas subjects of all other verbs will be marked differently, thus forming a less semantically-coherent ‘elsewhere’ class.
What, then, allows us to talk about ‘semantic alignment’ as a unitary phenomenon despite the observed cross-linguistic diversity of factors which determine it?

Proto-properties, although logically independent of each other, are cognitively intercorrelated: it is ‘natural’ or ‘unmarked’ for a participant volitionally involved in an event to cause it, and for a participant undergoing change of state to be affected by it; moreover, it is ‘natural’ for volitional participants to be involved in atelic activities and for affected participants to undergo telic changes of state (cf. inter alia Levin, Rappaport Hovav 1998, Ramchand 2003).

Unmarked constellations (volitional involvement in an activity, e.g. ‘walk’ and being affected by a change of state, e.g. ‘drown’) are cross-linguistically marked uniformly: their arguments are encoded as Agents and Patients, respectively (see Dowty 1991: 607).

Marked constellations (most crucially volitional involvement in a change of state, e.g. ‘sit down’, ‘start singing’ or ‘intentionally fall’ etc.) in different languages pattern differently depending on the relative weight individual languages give to particular semantic features: e.g., for Bats <volitionality> overrides <change-of-state>, whereas in Georgian it is vice versa.

Proto-property approach gives us a possibility to naturally account for non-binary ‘splits’ in argument encoding, cf. Choctaw (Muskogean; Gordon & Munro 1982: 84):

(44) a. čokma-li (45) b. sa-čokma (46) c. an-čokma
   good- 1SG.AG      1SG.PT-good      1SG.DAT-good
   ‘I act well’       ‘I am good’       ‘I feel good’

< A: volitional; causing >  < no >  < A: sentient; ( P : affected ) >

**Abbreviations**

ACT — active
AG — agentive
AOR — aorist
DAT — dative
DUR — durative
ERG — ergative
EXCL — exclusive
INCH — inchoative
NOM — nominative
PT — patientive
PRF — perfective
PRV — preverb
SB — subject agreement
ST — stative
References


