4

Thematic roles, event structure, and argument encoding in semantically aligned languages

PETER M. ARKADIEV

4.1 Introduction*

A major outcome of research in the realm of 'split intransitivity' or 'unaccusativity' in various theoretical frameworks of both 'functionalist' and 'formalist' kinds (see e.g. Perlmutter 1978, Rosen 1984, Merlan 1985, Van Valin 1990, Verhaar 1990, Mithun 1991, Levin & Rappaport Hovav 1995, 2000, Primus 1999, Alexiadou et al. 2004a) is the more or less general agreement that the phenomena usually subsumed under these terms are semantically driven. Moreover, some researchers (e.g. Merlan 1985, Van Valin 1990, Mithun 1991, Zaenen 1993) have pointed out that these phenomena and, more importantly, their determining factors show considerable variation, both cross-linguistic and intra-linguistic.

The semantic parameters which have most prominently figured in the discussion of 'semantic alignment' are the following two:

- the thematic role of the predicate's participant, especially its position on the Agent–Patient continuum;
- the inherent aspect of the predicate, especially the dichotomy between 'stative' and 'dynamic' predicates and between 'telic' and 'atelic' predicates.

In different languages one of the two parameters may become more prominent; thus, Mithun (1991) claims that in Guaraní stativity is the determining factor, while in Lakota it is agentivity.

* I am grateful to Bernard Comrie, Greville G. Corbett, Mark Donohue, Dmitry Gerasimov, Alexander E. Kibrik, Andrej A. Kibrik, Marian Klamer, Leonid I. Kulikov, Andrej L. Malchukov, Marianne Mithun, Andrej Nefedov, Barbara H. Partee, Andrej Shluinsky, Yakov G. Testelets, and Søren Wichmann for valuable comments on talks on which this paper is based, on the paper itself, and for other kinds of help. All faults are mine. I express my deep gratitude to the Max Planck Foundation, whose sponsorship made it possible for me to take part in the Conference on the Typology of Active-Stative Languages in Leipzig. This research was also partially funded by the Russian Science Support Foundation and by the Section of History and Philology of the Russian Academy of Sciences.

The major question still remaining on the agenda of current research and discussion in the field concerns the universality of proposed parameters and the problem of their interrelatedness. Indeed, in-depth studies of semantic alignment phenomena in particular languages often suggest that the factors determining argument encoding are language-specific, although typologically stable, and that it is usually a single factor or a group of closely connected factors which plays the determining role in a given language. However, it would be more desirable to have a typologically coherent picture of the phenomenon of semantic alignment and of the semantic parameters which underlie it.

Recently, Primus (1999) has proposed a universalist account of semantically aligned argument encoding with single-argument (hereafter monadic) predicates, following Dowty (1991), who proposed the well-known decomposition of thematic relations into separate 'proto-properties', which verbs entail with respect to their arguments. First of all, let us review the properties of prototypical Agent and Patient introduced by Dowty (1991) (see (1) and (2)).

(1) 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).

(2) 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: 572–4) but significantly intercorrelated (cf. e.g. Hopper and Thompson 1980 for correlations between various semantic parameters of transitivity); this intercorrelation is of utmost importance (see section 4.3).

The essence of Primus's proposal is her Principle of Morphosyntactic Expression of Thematic Information (Primus 1999: 90–100; cf. also Kibrik 1997):

(3) The more Proto-Agent (resp. Proto-Patient) properties an intransitive (in our terms monadic) 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 (in our terms dyadic) predicate.

What is important about this proposal is the fact that it requires 'counting' and 'balancing' of proto-properties in order to determine the type of encoding the verb imposes upon its sole argument. This proposal, undoubtedly, is conceptually appealing, especially for those linguists who insist on the prototypical organization of grammatical and lexical categories (e.g. Croft 1991, 2001). However, as I shall show in subsequent sections, it fails to predict the actual distribution of

intransitive argument-encoding types in particular languages. It turns out that it is usually possible to discern a single proto-property which determines the 'split' between Agent-like and Patient-like encoding (cf. similar observations in a recent illuminating discussion of semantic alignment phenomena in Austronesian languages in Foley 2005).

In the next section I will survey the data from several languages, most of which have featured prominently in the discussions of semantically aligned argument encoding, and try to show that in each of these languages it is possible to pinpoint a single factor, i.e. a unique proto-property or a combination of proto-properties, which determines the encoding of the single argument of monadic verbs as either Agent-like or Patient-like. More importantly, these prominent factors, which become grammaticalized in particular languages, turn out to be 'stronger' then all other proto-properties regardless of their number. That is, if a verb entails for its argument a specific property α from, say, the set of agentive entailments, then this argument gets A-like coding even if this verb entails with respect to it a greater number of patientive properties.

4.2 Case studies

4.2.1 *Loma: a genuine 'active-stative' language*

Loma (or Looma) is a Southwestern Mande language spoken in Liberia and Guinea.[1] As in other Mande languages, Loma has a system of several 'series' of pronouns which, besides encoding person and number, encode also such genuinely verbal categories as tense, aspect, mood, as well as polarity. Among these sets of pronouns there are two whose distribution is determined by the grammatical function of the relevant participant: one set ('Subjective') is used for As of dyadic predicates, the other ('Objective') encodes Ps of dyadic predicates. This is shown in Table 4.1 and examples (4) and (5).

(4) *gá té γa zu*
1PL.EXCL.SBJV 3PL.OBJV see DUR
'We see them.'

(5) *gà ká zu*
1SG.SBJV see+3SG.OBJV DUR
'I see him.'

If we turn to monadic predicates, we find that most of them require 'Subjective' encoding of their sole argument, identical to the A of dyadic predicates (see (6) and (7)); however, there is a class of predicates, namely stative verbs, whose sole argument is encoded as the P of dyadic predicates (see (8)).

(6) *gá li zu*
1PL.EXCL.SBJV go DUR
'We are going.'

(7) *tóa lo zo*
3SG.SBJV fall DUR
'He is falling.'

[1] My primary source of data is an article by Rude (1983), which is specifically dedicated to non-nominative patterns of alignment found in this language. All examples come from this paper; see Vydrin (1987) for a detailed description of the language.

Table 4.1. 'Subjective' and 'objective' pronouns in Loma

Person	'Subjective'		'Objective'	
	Singular	Plural	Singular	Plural
1st	*gà*	*gá* (excl.) *dá* (incl.)	low tone[a]	*gé* (excl.) *dé* (incl.)
2nd	*jà*	*wá*	*è*	*ù*
3rd	*tóa*	*tá*	high tone[a]	*té*

[a] In these cases tonal features are attached to the verbal stem, cf. (5), where the verbal stem *ka* 'see' is combined with the high-tone autosegmental morpheme expressing 3rd person singular; voicing of the verbal stem's initial consonant, cf. (4), always occurs before an overt Objective pronoun.

(8) *gé* *βala βɛ*
1PL.EXCL.OBJV big DUR
'We are big.'

Thus S encoding in Loma is determined by the verb's inherent aspect, i.e. the stative ~ dynamic opposition, as represented in (9):

(9) <Pred: Stative> → <S: Objective>
<Pred: Dynamic> → <S: Subjective>

It is important to note that some dynamic predicates, such as 'fall' in (7), entail for their sole argument more patientive proto-properties (e.g. 'change of state', 'affected') then agentive, and would be predicted by Primus's generalization (3) to require 'Objective' rather then 'Subjective' encoding of their arguments. However, this is not the case: in Loma, if a verb entails a purely patientive proto-property 'undergoes change of state' for its sole argument, the latter is encoded as A and not as P.

It is interesting that stative predicates formed from nouns with the help of the copular verb *gà* behave like dynamic verbs and encode their S argument with the 'Subjective' set of pronouns (see (10)):

(10) *tá* *gà* *zunu*
3PL.SBJV COP man
'They are men.'

It is necessary to note that the distribution of 'Subjective' and 'Objective' sets of pronouns outlined above is observed in Loma only in the imperfective aspect; once the aspect is switched to perfective, all distinctions are neutralized in favour of the 'Objective' set, which now encodes not only all types of S, but also As and Ps of dyadic verbs (see (11) and (12)).

Table 4.2. Verb classes in Georgian

Class	Case marking in Series II	Agreement (3sg Present, 3pl Present, 3pl Aorist)
I	A: Ergative—P: Nominative	*-s/-en/-es*
II	S: Nominative	*-a/-ian/-nen*
III	S: Ergative	*-s/-en/-es*

(11) *té gé γa ne*
3pl.objv 1pl.excl.objv see pfv
'They saw us.'

(12) *gé li ni*
1pl.excl.objv go pfv
'We went away.'

To conclude, Loma constitutes a rather rare example of a language where (in a subset of clause types at least) the choice of S encoding is determined solely on the basis of the aspectual distinction 'stative' vs. 'dynamic'.

4.2.2 *Georgian: telicity*

In Georgian, a Kartvelian language of the southern Caucasus, semantic alignment manifests itself in the partitioning of the verbal lexicon into several classes (see all the particulars and representative lists of verbs in Vogt 1971 and Harris 1981). There are four classes, but only three of them are really productive; verbs belonging to these classes show different subject agreement suffixes and, more importantly, different case marking of arguments in Aorist and Optative (the so-called Series II of tenses)[2] (see Tables 4.2 and 4.3).

Dyadic verbs belong to class I, which mark their A participant with Ergative, and their P participant with Nominative, while monadic verbs are distributed among the two other classes. According to Holisky (1979, 1981) and Harris (1981, 1982; cf. also Van Valin 1990), the principal rationale behind the assignment of monadic verbs either to class II or to class III lies in the realm of telicity: most telic verbs, denoting change of state, fall into class II, while most atelic verbs,

Table 4.3. Example verb forms of the three classes

	Class I 'paint'	Class II 'die'	Class III 'work'
3sg Present	*xat'av-s*	*k'vdeb-a*	*mušaob-s*
3pl Present	*xat'av-en*	*k'vdeb-ian*	*mušaob-en*
3pl Aorist	*daxat'-es*	*mok'vd-nen*	*imušav-es*

[2] The same distinctions manifest themselves also in Perfect and Pluperfect (the so-called Series III of tenses).

denoting unbounded processes and activities, fall into class III. It is important to note that agentivity proper does not play a substantial role in the system of Georgian verb classes; indeed, both classes II and III comprise verbs whose S arguments may have various degrees of agentivity. Consider first 'prototypical' examples of agentive atelic verbs in (13) and (14), and patientive telic verbs in (15) and (16):[3]

(13) *k'ac-ma i-lap'arak'-a*
man-ERG PFV-talk-AOR.3SG.SBJV
'The man talked.'

(14) *gogo-eb-ma i-tamaš-es*
girl-PL-ERG PFV-play-AOR.3PL.SBJV
'The girls played.'

(15) *c'q'al-i ga-tb-a*
water-NOM PRV-warm.up-AOR.3SG.SBJV
'The water became warm.'

(16) *k'ac-i mo-k'vd-a*
man-NOM PRV-die-AOR.3SG.SBJV
'The man died.'

Verbs such as 'talk' or 'play' entail for their S participant quite a lot of agentive properties, such as 'volition', 'causing', 'sentience', often 'movement', and no patientive properties. On the contrary, verbs such as 'warm up' or 'die' entail for their sole argument only patientive properties, such as 'affected', 'change of state', 'stationary', often 'incremental theme'. On the basis of these examples only, we could argue that class II verbs are those which entail enough patientive properties with respect to their S argument, while class III verbs are those which entail for it enough agentive properties. However, this is not the case, as examples such as (17–19) show.

(17) *bavšv-eb-i da-sxd-nen*
child-PL-NOM PRV-sit.down-AOR.3PL.SBJV
'The children sat down.'

(18) *k'ac-i a-mγer-d-a*
man-NOM INCH-sing-INCH-AOR.3SG.SBJV
'The man began to sing.'

(19) *c'q'al-ma i-duγ-a*
water-ERG PFV-boil-3SG.SBJV
'The water boiled (for some time).'

Examples (17) and (18) show class II verbs whose S arguments are clearly agentive; these verbs actually entail for their sole argument more agentive proto-properties

[3] All examples come from Nino Amiridze and Ketevan Gadilia, whom I thank for generous help.

('volition', 'sentience', 'causing', 'movement') than patientive ('change of state' only). However, it is the latter property which serves as the sole determinant of class assignment. On the other hand, the verb 'boil' in (19) shows some clear patientive properties ('affected', 'stationary'), and probably no agentive ones; however, it falls into the same class as 'talk' or 'play'. What is crucial is the absence of the 'change of state' entailment for non-agentive verbs of class III (most of which denote processes which involve inanimate entities and have a salient observable outcome, such as 'shine', 'glitter', 'roar (of water)': see Holisky 1981 for a comprehensive account).

To summarize, I propose the following generalization about verb class assignment for monadic verbs and case marking in Georgian (see (20)):

(20) < θ: {P: 'change of state'}> → <Pred: Class II>, <S: Nominative>
< θ: elsewhere> → <Pred: Class III>, <S: Ergative>

The only notable exception to (20) is the behaviour of stative predicates formed with the aid of the copula; these invariably pattern with class II verbs, allowing only Nominative encoding of their S argument: cf. (21) (Harris 1981: 250).

(21) *tamar-i* *iq'o* *mepe.*
Tamar-NOM COP.AOR.3SG monarch-NOM
'Tamar was the monarch.'

Thus, just as in Loma, copular predicates are exceptional to a semantically driven rule.

4.2.3 *Bats and Tabassaran: volitionality*

Bats (or Tsova-Tush) and Tabassaran, two Nakh-Dagestanian languages of the north Caucasus, show rather similar patterns of semantic alignment, differing mainly in the degree of 'fluidity'. All examples from Bats come from Holisky (1987); Tabassaran data is taken from Kibrik (1985).

Bats shows 'fluid' semantic alignment only in pronominal S participants of monadic verbs: cf. (22a) and (22b), where the same verb 'fall' allows for both Nominative and Ergative marking of S with a clear difference in volitionality:

(22) a. *as* *wože*
1SG.ERG fell
'I fell (on purpose).'

b. *so* *wože*
1SG.NOM FELL
'I fell (by accident).'

There are three main classes of monadic verbs in Bats (more representative lists in Holisky 1987: 122–30):

1. those which take only Nominative Ss, e.g. *maicdar* 'be hungry', *qerl'ar* 'be afraid', *dah"γordar* 'freeze', *daq'dalar* 'dry up', *daxdalar* 'go mindlessly, unconsciously', *mildar* 'be cold', *q'eč'ar* 'to appear (of parts of objects)';

2. those which take only Ergative Ss, e.g. *dat'ar* 'run', *daxar* 'go', *eq:ar* 'jump', *dadar* 'swear', *axar* 'bark', *lavar* 'talk', *darc'dalar* 'take off one's clothes', *da:xar* 'live', *lap'c'ar* 'play';
3. those which allow both Nominative and Ergative Ss, e.g. *dah"daxar* 'get drunk', *derc'ar* 'turn into', *xarcdalar* 'change', *ʕopdalar* 'hide, come out of sight', *k'určdalar* 'roll', *dožar* 'fall over', *qac'ar* 'be in a hanging position'.

Verbs taking only Nominative Ss denote states and events whose participant, even if human, is clearly non-volitional; that this lack of volitionality is a lexicalized property of these verbs is clearly indicated by the lexeme *daxdalar* 'go unconsciously'. On the contrary, verbs with Ergative Ss only denote activities performed by volitional human Agents, consciously aware of their actions. Finally, those verbs which allow both Nominative and Ergative marking of S (they also fall into several subtypes according to the preferred pattern of argument encoding) denote situations whose participant may be either volitional (hence Ergative marking) or non-volitional (hence Nominative marking). For instance, the verb *ʕopdalar* means 'subject comes to be hidden not because of anything (s)he herself does, but because something moves in front of her/him' with Nominative S, and 'subject does something which results in her/him becoming hidden, e.g. moves behind a barrier' with Ergative S.

There are some peculiar cases of shifts in case marking with verbs for which only one interpretation (viz. volitional or non-volitional) is pragmatically neutral and 'unmarked'. For example, the prototypically non-agentive verb *dah"davar* 'die' with Ergative S means 'to die because of either doing something in order to die or doing nothing to prevent death', whereas the prototypically agentive verb *ga=rek'adalar* 'run' with Nominative S is interpreted as 'to run unwillingly, e.g. because of a very steep path'. Also, the verb *kebadalar* with Ergative S means 'to boast', and with Nominative S 'to be praised'. What is important for the present discussion is that shift of argument encoding is sensitive to the single feature 'volitional', which may override the whole array of proto-patientive properties, as with the verb 'to die', or whose lack is more important than the presence of other proto-agentive entailments, as with the verb 'to run'. These observations are formally represented in (23):

(23) < θ: {A: volitionality}> → <S: Ergative>
< θ: elsewhere> → <S: Nominative>

We see a similar picture in Tabassaran. In this language there are two main sets of agreement suffixes on the verbs; with dyadic verbs one set ('Agentive') refers to the A participant, and another ('Patientive') to the P participant: see Table 4.4 and example (24).

(24) *uzu uvu* ʀ*urč*$_w$*un-za-vu*
1SG 2SG beat-1SG.AGT-2SG.PAT
'I have beaten you.'

Table 4.4. Agentive and Patientive agreement in Tabassaran

Person	Agentive		Patientive	
	Singular	Plural	Singular	Plural
1st	*-za*	*-a*	*-zu*	*-u*
2nd	*-va*	*-c_wa*	*-vu*	*-c_wu*

The situation with monadic verbs is similar to that found in Bats: there are three main classes: (i) verbs which take only Agentive agreement, e.g. *daqun-za* 'I lay down', *ʀižun-za* 'I began to cry', *ʀušun-za* 'I came'; (ii) verbs which take only Patientive agreement, e.g. *ǩabqun-zu* 'I drowned', *ʀarʀun-zu* 'I froze', *ergra-zu* 'I got tired'; and (iii) verbs which may take both Agentive and Patientive agreement, e.g. *ʀuʒun-za* 'I remained (voluntarily)' vs. *ʀuʒun-zu* 'I remained (against my will)'; *aqun-za* 'I fell (intentionally)' vs. *aqun-zu* 'I fell (by accident)'; *hilirqun-za* 'I shook (on purpose)' vs. *hilirqun-zu* 'I shook (involuntarily)'. Verbs with Agentive agreement only denote volitional and controlled events, while those which allow only Patientive agreement denote non-volitional events. The class of verbs unspecified for the volitional entailment is smaller in Tabassaran than in Bats, and the process of 'recategorization' of volitional or non-volitional verbs 'by default' into the opposite classes is here less productive. Thus, we can see that languages may differ as to the degree of conventionalization and lexicalization of particular proto-properties.

4.2.4 *Central Pomo: affectedness*

Let us now turn to a more complicated case, that of pronominal case marking in Central Pomo, a Pomoan language of California, as described by Marianne Mithun (1991, 1999: 217–19, this volume).[4] There are three cases, Nominative, Accusative, and Oblique; I will focus only on the first two.

With dyadic verbs, the A participant is encoded by Nominative case and the P participant by the Accusative case: see examples (25a) and (25b):

(25) a. *ʔa:* *mú:-ɬu* *ʔé:čadiw*
1sg.nom 3sg-acc chase.away
'I chased him away.'

b. *mu:-l* *ɬo:* *ʔé:čadiw*
3sg-nom 1sg.acc chase.away
'He chased me away.'

The situation with monadic verbs is rather complex. As in Bats and Tabassaran, there are three main classes: verbs which consistently require Nominative encoding of the S, verbs which allow only Accusative encoding of the S, and those which

[4] Similar patterns are found also in other languages of this family; see e.g. O'Connor and Caissee (1981) for Northern Pomo.

allow both. However, the distribution of lexemes into these three morphosyntactic classes follows a system quite different from that observed in the languages of the Caucasus.

Verbs taking only Nominative Ss may be dynamic, cf. (26), or stative, cf. (27):

(26) a. *ʔa:* *wáq'iʔle*
1SG.NOM would.go
'I would go.'

b. *ʔa:* *sbíč'*
1SG.NOM get.up
'I got up.'

c. *ʔa:* *pʰadé:n*
1SG.NOM swim
'I swam.'

d. *ʔa:* *čáč'*
1SG.NOM escape
'I escaped.'

(27) a. *ʔa:* *yá:qač'in*
1SG.NOM careful
'I'm careful.'

b. *ʔa:* *ʔnáč'*
1SG.NOM hiding
'I'm hiding.'

c. *ʔa:* *ʔe* *qʰám q'dí.*
1SG.NOM COP kindhearted
'I'm kindhearted.'

d. *ʔa:* *ʔe* *kú:č'*
1SG.NOM COP mean
'I'm mean.'

Dynamic verbs of this class are clearly agentive and imply volitionality on the part of their sole participant; however, it is not possible to say the same about the stative predicates with Nominative Ss: most of them denote states inherent to the participant and thus uncontrollable, but these states also do not significantly affect the individual of which they are predicated (note, however, that among these predicates many are formed with the use of the copula (cf. (27c), (27d)), and thus may perhaps be better accounted for as exceptional, like those in Loma and Georgian[5]).

Verbs with obligatory Accusative marking of the sole participant also include both dynamic and stative predicates. Dynamic predicates are exemplified in (28), and all of them imply both 'affectedness' and 'change of state', as well as lack of volitionality.

(28) a. *ɬo:* *ló:ya*
1SG.ACC fall
'I fell.'

b. *ɬo:* *madáts'čiw*
1SG.ACC slip
'I slipped.'

c. *ɬo:* *škʰé:nada.*
1SG.ACC get.well
'I'm getting well.'

d. *ɬo:* *qamáʔleč*
1SG.ACC get.angry
'I got angry.'

[5] However, a similar 'split' between inherent and transient states, where the former encode their sole argument as A and the latter as P, is found also in the Austronesian language Larike (Klamer, this volume), where both types of predicate are uncontroversial verbs. This shows that the distribution found in Central Pomo is probably not accidental.

Stative predicates of this class are shown in (29).

(29) a. *ɬo:* *kasíla*
1SG.ACC cold
'I'm cold.'

b. *ɬo:* *kíts'čiw*
1SG.ACC scared
'I'm scared.'

c. *ɬo:* *ʔtʰál*
1SG.ACC painful
'I'm in pain.'

d. *ɬo:* *mká:t'*
1SG.ACC surprised
'I'm surprised.'

Examples in (29) are of particular interest, since they are characterized by a set of entailments coming from both Patientive set ('affected') and Agentive set ('sentient'): these are typically temporary states of an animate being which is capable of perceiving its being in such a state.

Finally, verbs allowing for both Nominative and Accusative marking of the S participant are mainly dynamic, and the difference in case marking is driven by volitionality (cf. examples 30 and 31 from Mithun 1999: 218).

(30) a. *ʔa:* *sma* *mtí:č'*
I.NOM go.to.bed
'I went to bed.'

b. *ɬo:* *sma* *mtí:č-ka*
I.ACC go.to.bed-INFER
'I must have fallen asleep.'

(31) a. *ʔa:* *k'lú:k'lu:w*
I.NOM cough
'I coughed (intentionally).'

b. *ɬo:* *k'lú:k'lu:w*
I.ACC cough
'I coughed (accidentally).'

In order to adequately characterize the argument-marking patterns found in Central Pomo, one has to posit three rules instead of two, and stipulate an ordering between them, as in (32).

(32) a. <θ: {A: volitional}> → <S: Nominative>
b. <θ: {P: affected; A: sentient}> → <S: Accusative>
c. <θ: elsewhere> → <S: Nominative>

Indeed, it seems that although Accusative marking of Ss is clearly a marked option in Central Pomo (cf. Mithun's observation that a participant has to be significantly affected in order to count for a 'real' Patient in this language), and thus requires an explicit statement in terms of the proto-property 'affected' (32b), Nominative marking for some types of dynamic verb, among which are those which do not imply significant affectedness of the S (such as *maɬ'ém* 'to step on something'), is evidently determined by the presence of the volitional entailment (32a). The fact that Nominative marking is nevertheless default for Ss in Central Pomo is captured by the elsewhere rule (32c).

4.2.5 *Summary*

The data from five languages I have presented in this section clearly indicates that languages do not mark the sole argument of monadic verbs on the basis of the overall balance of agentive and patientive proto-properties, as Primus (1999)

argued. Rather, each language selects a single property or a group of properties which determines argument marking, and in each language this privileged entailment overrides all entailments from the opposite set. Thus, in Georgian, agentive telic verbs (e.g. 'sit down') fall into one class with patientive ones (e.g. 'die') on the basis of the 'change of state' entailment, whereas in Bats it is possible to construe such clearly patientive predicates as 'die' as involving volition on the part of the S argument. Finally, in Central Pomo the determining criterion for patienthood is not a single property from the patientive set of entailments, and not even a group of patientive entailments, but rather a combination of a patientive entailment ('affectedness') with an agentive one ('sentience').[6]

Therefore, the most obvious conclusion one may draw from the previous discussion is that, although Dowtyan proto-properties are certainly a useful means of describing the phenomenon of semantic alignment (cf. Foley 2005), such straightforward applications of them as Primus's universalist proposal are empirically inadequate. In the next section I discuss some theoretical implications of this fact in more detail.

4.3 Theoretical implications

The discussion of the data in the previous section has shown that in order to capture the actual distribution of argument-marking patterns in semantically aligned languages, one has to pinpoint a single factor or group of factors (statable in terms of thematic proto-properties) rather then calculate an overall balance of agentive and patientive entailments of predicates. Thus, it may seem that we arrive at the very conclusion from which we started: that in different languages different parameters are responsible for argument encoding with monadic verbs, and that these parameters are neither interdependent nor allow for a uniform typological characterization. However, I would like to show that this is not the case: in spite of clearly observable cross-linguistic diversity, semantic parameters of semantic alignment are intrinsically intercorrelated and show a typologically coherent basis.

Let us first consider two related cross-linguistic facts.

First of all, it is possible to observe that despite all diversity there are 'prototypical' agentive (or patientive) predicates which in all languages with semantic alignment tend invariably to encode their S arguments like As (or Ps) of dyadic verbs. If we disregard Loma—which is not a 'typical' semantically aligned language, since the 'split' in argument marking is here conditioned solely by the aspectual property 'stativity'—it turns out that those predicates which denote volitionally performed atelic activities (such as 'run', 'play', 'work') tend always to require A-encoding (unless, as in Bats and other languages with 'fluid' alignment—e.g.

[6] The entailment 'sentient' may even be grammaticalized on its own in some languages, e.g. in the Muskogean languages Choktaw and Chickasaw (Gordon and Munro 1982), where a tripartite rather than a binary 'split' is observed with monadic predicates. However, these issues are so complex that they require a separate cross-linguistic study.

Chol, see Gutiérrez and Zavala Maldonado 2005—the volitionality entailment is cancelled), while those that denote uncontrolled telic events (such as 'die', 'fall', 'drown') more often then not demand P-encoding of S (again, languages with 'fluid' alignment may show some variability here[7]).

Second, on the contrary, it is precisely those verbs which entail non-homogeneous sets of proto-role entailments for their sole participants that show most cross-linguistic variation. Indeed, such verbs as 'sit' and 'get up' belong to the 'patientive' class in Georgian and Chol (Gutiérrez and Zavala Maldonado 2005), but to 'agentive' in Bats and Central Pomo, while such verbs as 'boil' or 'tremble' show opposite patterning: they are clearly 'patientive' in Bats but 'agentive' in Georgian. Similar discrepancies between languages are often cited in the literature (see Rosen 1984, Merlan 1985, Sorace 2000[8]).

These two facts point towards the following conception of the mutual relation of semantic parameters of semantic alignment: factors pertaining to different proto-properties, most importantly to such features as [±volitionality], [±affectedness], [±change of state], are, indeed, logically independent of each other, as shown already in Dowty (1991), just like the different transitivity parameters proposed by Hopper and Thompson (1980); however, these features tend to be intercorrelated in that some combinations of features or entailments are conceived as more 'cognitively marked' (using Kibrik's 2003 term) and less 'natural' then others. It is not logically necessary for all monadic predicates which entail [+volitionality] to entail [–change of state] or vice versa, but it is significant that most monadic predicates with volitional S arguments are atelic, and most predicates which denote change of state entail affectedness and lack of volitionality. This observation is corroborated by the fact that with dyadic predicates these features are usually distributed between different participants: it is the volitional and controlling Agent who performs a certain activity, as a result of which a non-volitional affected Patient undergoes a change of state (see e.g. Foley and Van Valin 1984, Levin and Rappaport Hovav 1998, Testelec 1998).

Thus, we arrive at the following generalization:

(33) a. Unmarked combinations of features require consistent argument marking:
$<\theta$ {A: +volitional}, {P: –change of state}$>\rightarrow$ Agentive marking;
$<\theta$ {A: –volitional}, {P: +change of state}$>\rightarrow$ Patientive marking.

[7] An interesting counterexample comes from the aforementioned Choktaw, where, as claimed by Rosen (1984: 61), the verb 'to die' requires pronominal prefixes from the 'agentive' set; this may be due to the obligatory animacy of the S participant of this verb.

[8] The latter paper is an in-depth study of auxiliary selection with monadic verbs in four European languages; this phenomenon, as observed already by Rosen (1984) and Van Valin (1990), is clearly motivated by the same range of factors as argument encoding in semantically aligned languages.

b. Marked combinations of features are resolved by language-specific ranking of entailments:
$<\theta$ {A: +volitional}, {P: +change of state}$>\rightarrow$ Agentive marking if [±volitionality] >> [±change of state];
$<\theta$ {A: +volitional}, {P: +change of state}$>\rightarrow$ Patientive marking if [±change of state] >> [±volitionality].

In Georgian, where telicity ([±change of state]) is more prominent then agentivity ([±volitionality]), agentive telic predicates such as 'sit down' or even ingressive predicates such as 'begin to sing' or 'start working' require their S participants to be encoded like Ps of dyadic verbs. In Bats, on the contrary, where agentivity outranks telicity, such predicates encode their sole arguments as As of dyadic verbs.

Certainly, languages tend to conceive of such properties as 'change of state' and 'volitionality' as gradual rather then binary, all-or-none distinctions, and there is a certain degree of cross-linguistic variation with respect to where precisely the boundary between telic and atelic or agentive and non-agentive predicates is drawn. However, the generalization in (33) seems to be a reliable constraint on typological variation in the realm of semantic alignment (cf. similar proposals by Foley 2005 based on a careful survey of data from Austronesian languages).

4.4 A perspective from event structure and 'unaccusativity'

For most of the three-decade period of intensive studies of the range of phenomena subsumed under the term 'semantic alignment' this research was pursued independently of, and without much interest in, parallel studies in the realm of so-called 'unaccusativity', although the basic similarity of the semantic factors underlying both types of phenomenon was evident for the earliest students of 'unaccusativity' (see e.g. Perlmutter 1978 and Rosen 1984). Notable exceptions of attempts at unifying these two perspectives may be found in Verhaar (1990) and Van Valin (1990); see also Donohue (this volume). In this section I am going to try to investigate how recent developments in the study of 'unaccusativity', in particular work by Levin and Rappaport Hovav (1995, 2000), may be made useful for the discussion of semantic alignment.

An in-depth study of such phenomena as formation of resultative constructions and causatives in English and some other languages led Levin and Rappaport Hovav to the proposal that argument linking (assignment of deep syntactic roles Subject and Object, which may be more or less equated with e.g. Actors and Undergoers in Role and Reference Grammar: see Foley and Van Valin 1984, Van Valin 1990) is determined by the lexical semantics of the verb (it is important to note here that a particular lexeme, e.g. *run*, may have different semantic and, consequently, syntactic properties in different contexts: see Levin and Rappaport

Hovav 1995: ch. 5, 1998). Argument linking is implemented via the following linking rules (Levin and Rappaport Hovav 2000: 285–94):

(34) a. Immediate Cause Linking Rule
The argument of the verb that denotes the immediate cause of the eventuality described by that verb is its external argument [Actor—*PA*]

b. Directed Change Linking Rule
The argument of the verb that corresponds to the entity undergoing the directed change described by the verb is its direct internal argument [Undergoer—*PA*]

In order to account for such verbs that are neither telic (do not denote change of state) nor agentive, e.g. verbs denoting states—which in the languages surveyed by Levin and Rappaport Hovav behave like 'unaccusative' verbs, i.e. have an Undergoer argument—they propose also the following 'elsewhere' linking rule:

(34) c. Default Linking Rule
An argument that does not fall under the scope of the other two linking rules is a direct internal argument.

Finally, since the evidence of resultative constructions in English points toward the classification of agentive telic verbs such as verbs of directed motion as 'unaccusative', in cases when there is a conflict between the Immediate Cause and Directed Change linking rules, the latter is ranked above the former.

Such, in outline, is the theory of Levin and Rappaport Hovav. How could this theory be applied to the phenomena of argument encoding in semantically aligned languages? My claim is that the linking theory of Levin and Rappaport Hovav is to a certain degree equivalent to the Proto-Role theory outlined in the previous section; what it lacks is a greater degree of flexibility, which would account for the actual cross-linguistic variation in the realm of semantic alignment. Let us see now what the minimal possible amendments needed by the linking theory are.

If we consider first Georgian, we will find that the linking theory as it is more or less correctly predicts the distribution of verbs with Nominative and Ergative encoding of S: agentive atelic verbs select for an Ergative argument, telic verbs select for Nominative arguments regardless of agentivity (that means that in Georgian rule (34b) outranks (34a)). Non-agentive atelic verbs (at least those which are not stative), however, encode their S participant with Ergative rather than with Nominative, which means that their Ss are Actors rather then Undergoers in Georgian. This fact may be rather straightforwardly accounted for if we assume that Default linking rule in Georgian assigns participants which neither undergo directed change of state nor are agentive to the Actor argument, and not to the Undergoer, see (35):

(35) a. Default Linking Rule for Georgian
An argument that does not fall under the scope of the other two linking rules is an external argument (Actor).

Table 4.5. Argument linking in three semantically aligned languages

Language	Ranking of rules	Default rule
Georgian	Directed change > Immediate cause	External argument
Bats	Immediate cause > Directed change	Internal argument
Central Pomo	Immediate cause > Directed change	External argument

If we turn now to Bats and Central Pomo, we find that the original version of linking theory fails to predict the behaviour of agentive verbs of change of state such as 'sit down', as well as of genuinely patientive verbs such as 'die' which may be recategorized as agentive under certain pragmatic circumstances. Indeed, since rule (34b), pertaining to arguments undergoing change of state, is ranked before rule (34a), dealing with agentive arguments, there is no way to derive the actual case marking in these languages. What is needed, then, is simple reranking of rules: once (34a) is made higher in rank than (34b), the actual distribution of argument-encoding patterns in Bats, Tabassaran, and Central Pomo follows quite straightforwardly.

So, what we need to make the linking theory correctly predict the facts is again parameterization: we have to allow the linking rules to be differently ranked in different languages, and also different types of argument to serve as 'default'. Table 4.5 illustrates different 'parameter settings' for the languages surveyed in this chapter.[9]

Thus, a rather simple parameterization of linking rules allows the linking theory to correctly account for a broader range of phenomena of semantic alignment.

4.5 Conclusions

In this chapter I have shown that, despite the considerable cross-linguistic diversity found in the realm of phenomena subsumed under the term 'semantic alignment', it is possible to construct a coherent and conceptually rather simple theory which will account for both similarities and variation in this field. The principal semantic parameters of semantic alignment, which pertain to such fundamental notions as agentivity and telicity, can be unified as two facets of a more general phenomenon of 'event structure' (see Croft 1998, Levin and Rappaport Hovav 1998, Tatevosov 2002, Ramchand 2003). Though logically independent, they are nevertheless significantly intercorrelated, and languages tend invariably to encode 'unmarked' event structures, i.e. agentive activities and non-agentive changes of state. The

[9] The fourth logically possible type is exemplified by English and some other European languages, where the linking rules, however, do not affect the surface case marking of the S argument; whether there are semantically aligned languages of this type remains an empirical issue.

greatest degree of variation is predicted to be found in the domain of 'marked' event structures, especially agentive changes of state (see section 4.3).

The concept outlined in this chapter can be extended to a broader range of phenomena, viz. those subsumed under the heading 'unaccusativity' (see section 4.4). Whether all such phenomena in various languages are driven by the same or similar universal semantic factors is a question for further empirical investigation.